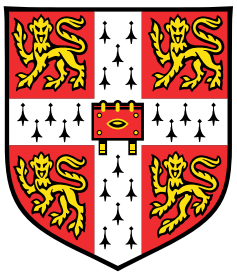
Accumulating delivered dose to the rectum to improve toxicity prediction in prostate radiotherapy

Doctoral Thesis



Leila Evelyn Aitken Shelley

Department of Engineering
University of Cambridge

This dissertation is submitted for the degree of
Doctor of Philosophy

Homerton College

July 2019

Declaration

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as specified in the text. It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as specified in the text. I further state that no substantial part of my thesis has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as specified in the text. It does not exceed the prescribed word limit for the relevant Degree Committee.

Leila Evelyn Aitken Shelley
July 2019

Abstract

Accumulating delivered dose to the rectum to improve toxicity prediction in prostate radiotherapy

Leila E. A. Shelley

Gastrointestinal (GI) toxicity is a clinical issue suffered by up to 22% of prostate cancer radiotherapy patients. However, the relationship between radiation dose and toxicity is generally poorly understood. In prostate radiotherapy, the rectum is a dose-limiting structure to which treatment planning dose constraints are applied to minimise the risk of toxicity. Current normal tissue complication probability (NTCP) models are based on planned dose data and do not consider the effects of organ motion on true delivered dose.

The VoxTox research programme has developed automated solutions for segmentation and dose calculation of the rectum for prostate cancer patients being treated with helical TomoTherapy. Daily image guidance scans, acquired primarily for the purposes of positional verification, are exploited by extracting quantitative information to facilitate the calculation of daily delivered and total accumulated dose to the rectum. Prospectively collected toxicity data at 2 years post-treatment were available for 295 patients across two separate cohorts.

In this thesis, the hypothesis being tested is that delivered dose is a better predictor of rectal toxicity than planned dose in prostate radiotherapy. The research has successfully demonstrated, for the first time, that delivered dose produces stronger associations with rectal bleeding and proctitis than planned dose. Analysis was performed using dose surface maps (DSMs) of the rectal wall, allowing spatial aspects of dose to be retained during accumulation. A subsequent analysis on a separate cohort also found stronger links between delivered dose and GI toxicity, stool frequency, and bowel bother, in addition to rectal bleeding and proctitis.

Biomechanical finite element (FE) modelling was introduced to provide a more anatomically plausible tool for dose accumulation and allowed more accurate tracking of dose at the voxel level. A sensitivity analysis was conducted which explored the effect of simulated rectal motion on dose, and corresponding change in NTCP. For VoxTox patient dose-toxicity analysis, further dose parameterisation approaches were explored in order to consider the

increased resolution of information available. Voxel-based rectal subregions at risk (SRRs) were identified using geometric and statistical approaches. In general, discriminative power improved with FE modelling for both planned and accumulated delivered dose, and associations between accumulated dose and toxicity were strengthened by voxel-based subregion analysis.

Multivariate NTCP models were constructed for 12 different toxicity endpoints based on planned and accumulated dose parameters. Model performance was compared between analysis approaches, and models were tested on a validation dataset. In general, FE-based dose models performed best, although the optimal dose parameter selected within the model varied with toxicity endpoint.

Overall, results suggest that there is an advantage to incorporating delivered dose into NTCP modelling. However, the differences between planned and accumulated dose can be subtle. Meaningful parameterisation of accumulated dose needs careful consideration, as traditional methods for quantifying planned dose may not be directly transferable. Voxel-based analysis techniques are recommended in order to accurately preserve and register spatial dose information, and have been shown to improve the strength of dose-toxicity associations. Further research into quantifying voxel level dose distributions is encouraged.

It is anticipated that the novel scientific contributions presented within this thesis will prove valuable for future development of clinical decision-making tools for adaptive radiotherapy, with the ultimate aim of reducing the incidence of radiation-induced toxicity for prostate cancer radiotherapy patients.

For Sam,
10 years on and I love you more than ever.
Thank you for believing in me unconditionally.
Now, onto our next adventure.

For Popop (1931 - 2016),
may we all strive to follow in your greatness.

For my godson Jacob (née 2017),
oh, the places you will go!

For Alf (née 2018),
our ray of golden sunshine.

Acknowledgements

I would like to thank my supervisors for their guidance and support throughout this PhD research: Prof Michael Sutcliffe, Cambridge University Department of Engineering; Dr Simon Thomas, Head of Medical Physics & Clinical Engineering, Addenbrooke's Hospital; Prof Neil Burnet, Principal Investigator of the CRUK VoxTox Research Programme until 2018, now Chair in Academic Proton Clinical Oncology at the University of Manchester.

The research conducted throughout this PhD project would not have been possible without the invaluable contributions of the multidisciplinary VoxTox team. Thank you to everyone, in particular Dr David Noble, Dr Marina Romanchikova, Karl Harrison, Amy Bates, Dr Raj Jena, Dr Jessica Scaife (in addition to those mentioned above).

I am very grateful for the support of the WD Armstrong Trust, whose contributions to promoting cross-disciplinary research between engineering and medicine allowed this PhD to transpire. I would like to acknowledge Cancer Research UK (CRUK) who provided the original funding for the VoxTox research programme. Thank you to Homerton College, the Institute of Physics and Engineering in Medicine, and the Institute of Physics for providing the funding that allowed me to attend many varied conferences and courses. Opportunities to present my work and learn from others have been invaluable and have provided many highlights during my PhD experience.

Thank you to my colleagues at the Edinburgh Cancer Centre for your flexibility and understanding. Particularly Linda Carruthers, Prof Bill Nailon, Michael Trainer, Dr Paul McGrane, and Jane Stewart. It's good to be back.

I am incredibly lucky to have my family as my greatest supporters. My rock, Sam, and our Goldendoodle, Alf; my mum (my hero), Joe, Nanny & our late Popop, Susannah, Dan, Robbie, Shona, Celeste, Poppy, Bernie; George, Angela, Jess, Sapphie; and to Dad & Mary – thank you for all of your support in making this happen. My incredible friends - NB, HBH, and beyond. Thank you also to Homerton College for many good times, particularly for introducing me to Tom, Sarah and little Lucy.

"If I have seen further than others, it is by standing on the shoulders of giants"
- Sir Isaac Newton

Table of contents

List of figures

List of tables

Chapter 1

Introduction

1.1 Radiotherapy for prostate cancer

Prostate cancer is the most common cancer amongst adult men in the UK. According to Cancer Research UK [32], 1 in 8 men will be diagnosed with prostate cancer in their lifetime. In 2015, there were approximately 130 new cases of prostate cancer diagnosed every day. By 2035, the incidence rates are projected to rise by 12 % to 233 cases per 100,000, but conversely, the rate of mortality is projected to fall by 16 % to 48 deaths per 100,000 [32, 100]. These projections suggests that increasingly more men are surviving their prostate cancer diagnosis. Ten-year survival rates are at 84 %, a significant improvement from the 1970s when just a quarter of patients survived beyond 10 years [32].

Radiotherapy is the primary course of treatment for approximately 30 % of prostate cancer patients [106]. Radiotherapy is the delivery of high-energy X-ray beams, typically 6-15 megavolts (MV), via a linear accelerator (linac) mounted on a ring or rotating gantry with the patient positioned at the central focus, or isocentre. Modern day rotational intensity-modulated radiotherapy (IMRT) involves the use of mechanical multi-leaf collimators (MLCs) to dynamically shape the radiation beam from multiple entry positions or continuous arcs, building up a high dose region tightly conformed to the tumour.

Delivered for both radical and palliative intent, radiotherapy can be prescribed as a stand-alone treatment, or to complement surgery and/or chemotherapy. The ionising properties of the radiation beam cause damage to the DNA double helix via double and single strand breaks resulting in chromosomal aberrations. Radiotherapy is delivered in a fractionated regime which exploits the difference in radiobiological response between tumour and healthy or 'normal' tissue by allowing time for normal tissue to repair whilst tumour damage accumulates. The therapeutic ratio is the relative probability of tumour cell kill to that of normal tissue complication [14].

The aim of any radiotherapy treatment is to maximise the dose to the tumour, whilst minimising dose to surrounding healthy organs. In prostate radiotherapy, the tumour volume is considered to be the prostate gland, which sits immediately anterior to the rectum and posterior to the bladder, as indicated by the contours in Figure 1.1a. As such, the rectum and bladder are considered to be dose-limiting organs at risk (OARs) which limit the total radiation dose that can be delivered to the prostate. Although tumour control improves with increasing dose, so does the incidence and severity of normal tissue damage. Thus, a balance is sought with sufficiently high dose for optimal probability of local control, whilst also maintaining an acceptably low level of incidence of severe normal tissue complications.

In 2016, the standard dose prescription in the UK changed from 74 Gy delivered in 2 Gy fractions over 37 fractions (7.5 weeks), to 60 Gy in 20 fractions of 3 Gy (4 weeks), as a result of the Conventional or Hypo-fractionated High Dose Intensity-Modulated Radiotherapy for Prostate Cancer 'CHHiP' trial [52, 108]. The 5-year outcome data demonstrated non-inferiority of the 60 Gy hypofractionated arm with respect to the 74 Gy conventional group, with comparable levels of late side-effects, or toxicity.

With increasing survival rates for prostate cancer patients, the consideration of post-treatment quality of life is becoming of greater importance in the clinical decision making process. The adverse effects of radiotherapy by can be debilitating and permanent [101], and yet, the mechanisms of radiation toxicity are poorly understood [14]. Radiation-induced rectal toxicity remains a clinical issue for up to a fifth of prostate cancer patients.

1.2 Radiotherapy treatment planning

In prostate radiotherapy, the prostate gland is treated as the gross tumour volume (GTV). A margin is grown around this volume to incorporate random and systematic error into the planning target volume (PTV), the volume to which the dose is prescribed. One example of random error is interfraction motion; variations in the day-to-day internal positioning due to anatomical or physiological movement. Interfraction motion affecting the position of the PTV can introduce deviations from the intended planned dose, so daily on-line image guidance is used prior to treatment delivery to ensure the prostate is in the intended treatment position.

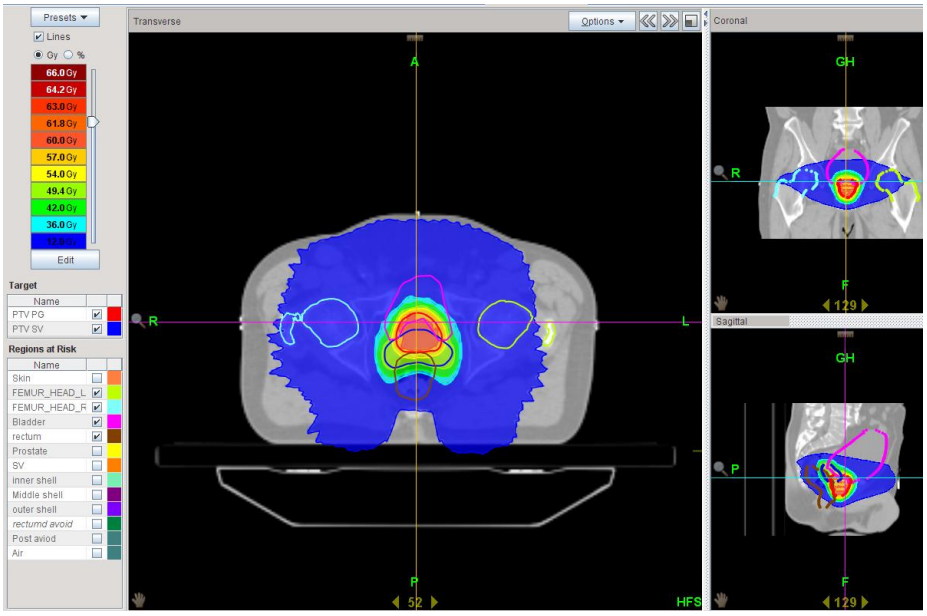
Each patient's radiotherapy treatment is individually planned based on a computed tomography (CT) scan of their anatomy. On this static CT image set, typically acquired a few weeks prior to the start of treatment, the PTV and any OARs are identified and contoured. Complex IMRT plans are generated using computational inverse-planning dose calculation algorithms within a treatment planning system (TPS). An example of a typical prostate

IMRT treatment plan is shown in Figure 1.1. Using the image guidance scans to register the prostate position at the time of treatment, with the reference planning CT scan, ensures that the radiation plan is being delivered to the PTV as intended.

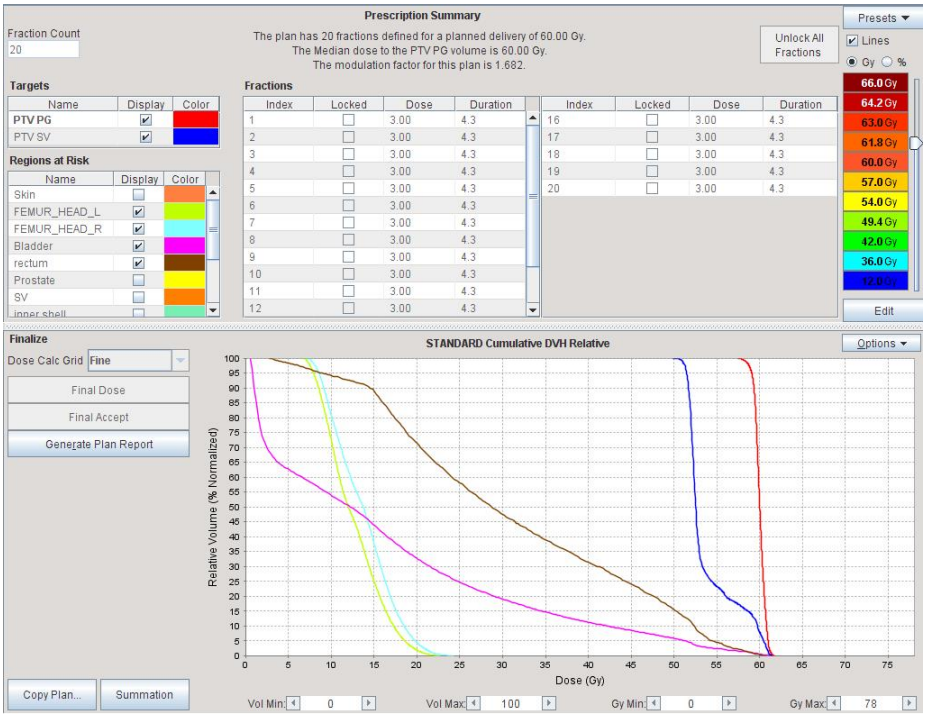
Exposing normal tissue within the treatment field to some level of concomitant dose is unavoidable due to the inherent nature of X-ray beam propagation through the patient. The OARs abutting the PTV are at risk of receiving dose levels associated with radiation-induced toxicity. Consequently, these OARs are dose-limiting factors when determining the radiation levels used to target the PTV. Treatment plans are optimised so that minimum dose coverage of the PTV is achieved, whilst also meeting OAR dose constraints. The OARs being considered in treatment planning for prostate radiotherapy are generally the rectum, bladder, and femoral heads. However, the bladder and rectum in particular undergo interfraction motion which is, aside from visually verifying IGRT scans for adherence with local setup protocols, generally unaccounted for.

Dose-volume constraints are organ-specific dose-threshold limits. For OARs, the upper threshold indicates the maximum acceptable level of 'Normal Tissue Complication Probability' (NTCP), below which, the risk of radiation-induced toxicity is reduced. In the case of the PTV, minimum thresholds are defined to ensure adequate target coverage for maximum 'Tumour Control Probability' (TCP). NTCP and TCP models in various forms have existed for several decades, historically derived from radiobiological models such as the Lyman-Kutcher-Burman (LKB) model [91, 86, 63, 28]. Practical implementation of NTCP dose-constraints in the clinic have been based upon published recommendations such as the 'Quantitative Analysis of Normal Tissue Effects in the Clinic' (QUANTEC) reports [94], as well as clinical trials. However, these models were generally derived based on 3D conformal radiotherapy and a thorough evidence base has not yet built up for complex radiotherapy techniques such as IMRT.

The dose received by a structure is commonly presented graphically using a dose-volume histogram (DVH); a plot of the cumulative dose-volume frequency distribution, as shown in Figure 1.1b. Dose-volume constraints are generally expressed as the percentage volume of a structure receiving less than or equal to a specified dose level. For example, the dashed black line in Figure 1.1b indicates that the dose received by 50% of the volume of the rectum is 28 Gy or less, i.e. $V_{50\%} < 28$ Gy. Existing DVH-based approaches to treatment planning have been criticised for lacking in spatial dose distribution information [3, 171]. With the standard of care now mainly high dose-gradient rotational IMRT (VMAT and helical TomoTherapy), geometric positioning and distribution of dose, as well as intra-organ sensitivity, become increasingly important considerations for dose-toxicity modelling.



(a)



(b)

Fig. 1.1 (a) Example of a TomoTherapy prostate radiotherapy treatment plan, prescribed 60 Gy in 20 fractions, with (b) corresponding Dose Volume Histogram (DVH). The plan shows axial, coronal and sagittal views centred about the prostate, with isodose colourwash from 12 Gy (20%) in blue, to 100% (60 Gy) in orange. Shown are prostate gland PTV (PTV PG, red) with seminal vesicle PTV (PTV SV, blue), sitting anterior to the rectum (brown), and posterior to the bladder (pink); right and left femoral heads are also contoured for dose reporting (cyan and pale green, respectively). The DVH is a cumulative dose-volume frequency distribution representing the absolute dose received by the percentage volume of the contoured structure.

The steep dose-gradients associated with IMRT mean that better understanding and tracking of interfraction motion of OARs is crucial. For organs such as the rectum in prostate radiotherapy, small deviations from the planned position can result in large differences in the true delivered dose. It has been shown that delivered dose to the rectum is not equal to planned dose, despite the use of image guided radiotherapy [139, 142]. The full extent of quantitative information available within image guidance scans has yet to be fully realised in routine clinical practice. Online image guidance scans can provide a platform for the calculation of delivered dose and facilitate a feedback system whereby dose can be monitored and accumulated, with remedial modifications made to the treatment plan as necessary [144, 172]. Ultimately, if delivered dose to OARs can be quantified, this should lead to improvements in NTCP modelling, and when considered in terms of adaptive radiotherapy, could result in reduced toxicity rates.

1.3 The VoxTox research programme

The VoxTox research programme is an observational study linking radiation dose to toxicity outcomes [46, 139]. The Cambridge-based project is a multidisciplinary collaboration spanning engineering, mathematics, high energy physics, medical physics and radiation oncology [30], which aims to combine sophisticated models of delivered dose with prospectively collected toxicity data, linking dose at the voxel level with toxicity.

Advanced imaging capabilities within the Radiotherapy Department at Addenbrooke's Hospital (Cambridge University Hospitals NHS Foundation Trust) provided the platform for calculation of delivered dose throughout treatment [30]. Megavoltage image guidance scans, acquired primarily for the purposes of tumour localisation, were used to calculate delivered dose by accounting for daily anatomical variation and provide a more accurate representation of the delivered radiotherapy treatment with respect to the static planning scan, as shown in Figure 1.2. Combining these data with prospectively collected clinical toxicity data allows us to address the hypothesis that delivered dose provides a better predictor of toxicity than planned dose.

VoxTox received approval from the National Research Ethics Service (NRES) Committee East of England (13/EE/0008) in February 2013 and is part of the UK Clinical Research Network Study Portfolio (UK CRN ID 13716).

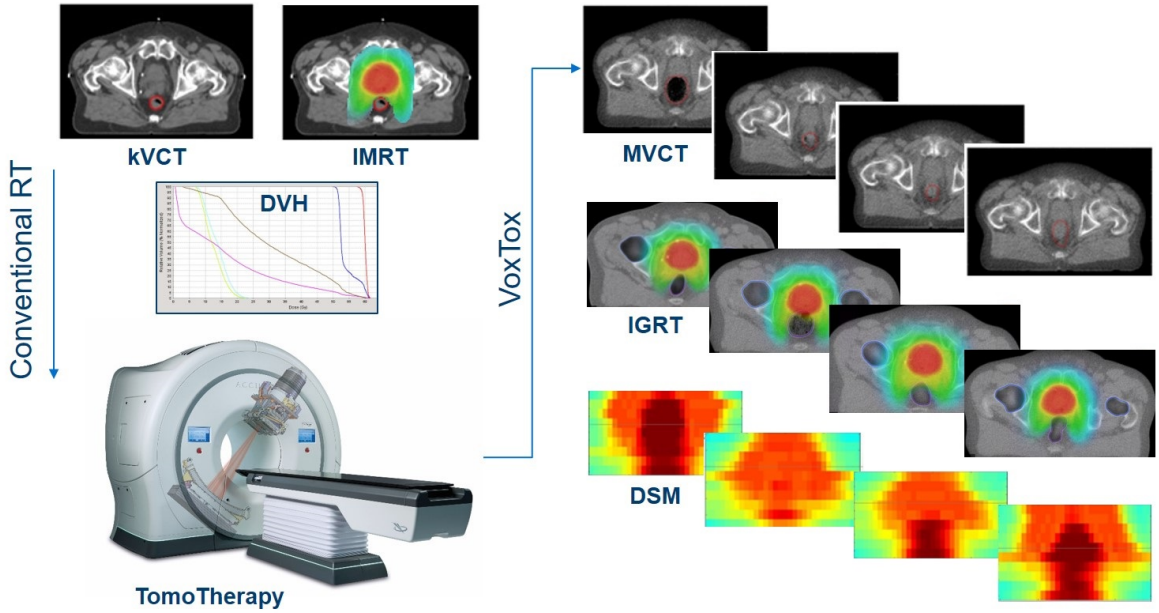


Fig. 1.2 VoxTox workflow. Images on the left show the conventional radiotherapy (RT) pathway, where anatomy is contoured on the kilovoltage computed tomography (kVCT) planning scan, allowing the intensity modulated radiotherapy (IMRT) plan to be generated. Plan quality is assessed using the dose volume histogram (DVH) parameters, and an optimal plan achieves adequate tumour coverage whilst minimising dose to healthy organs. On the right hand side, the VoxTox process mirrors this workflow on megavoltage computed tomography (MVCT) image guidance radiotherapy (IGRT) scans. Dose to the rectal wall is calculated and accumulated using dose surface maps (DSMs) which allow spatial features to be retained. Methods of parameterising DSMs are investigated within this thesis.

1.4 Research aims

The focus of this PhD project is to test the hypothesis of the VoxTox research programme, that delivered dose is a better predictor of toxicity than planed dose, specifically concentrating on rectal toxicity in prostate radiotherapy.

Delivered dose to the rectum and toxicity data collected for VoxTox prostate cancer patients treated between 2013-2018 have been collated and analysed. The analysis conducted within this thesis has depended upon the prior development of autosegmentation software, independent dose calculation algorithms, processing and tokenisation of patient data, finite element model development, and toxicity mapping systems, and was only possible due to the multidisciplinary nature of the VoxTox team.

Within the scope of this PhD research, delivered dose was initially represented using two-dimensional dose-surface maps of the virtually unfolded rectal wall. Biomechanical finite element modelling was later introduced as a more anatomically representative method for accumulating delivered dose. Parameterisation metrics were selected using modified versions of techniques previously applied in the literature for planned dose. These ranged from one dimensional reductions of the dose distribution, to the identification of toxicity-specific rectal subregions at risk. Additional information on the relationship between simulated rectal motion and resulting change in dose and NTCP were also investigated using biomechanical modelling. The research conducted in this thesis culminates in multivariate NTCP modelling developed on a training set for each of 12 toxicity endpoints, and tested on a validation set. Final models are presented based on parameters from both planned and accumulated delivered dose. A research overview is presented in Figure 1.3. It is anticipated that the outcomes of this PhD research will contribute towards the development of decision-making tools in adaptive radiotherapy, ultimately to reduce toxicity incidence for prostate cancer radiotherapy patients. A summary of the original contributions to knowledge are listed below:

- Demonstration that delivered dose to rectum can be more predictive of rectal toxicity than planned dose;
- Use of biomechanical finite element (FE) modelling to accumulate voxel-level delivered dose to rectum and link with toxicity;
- Biomechanical FE simulation of rectal motion to quantify dose and NTCP sensitivities;
- Voxelwise analysis of delivered dose to identify subregions associated with toxicity;
- The generation of rectal NTCP models, comparing multivariate models based on planned dose with those based on accumulated delivered dose.

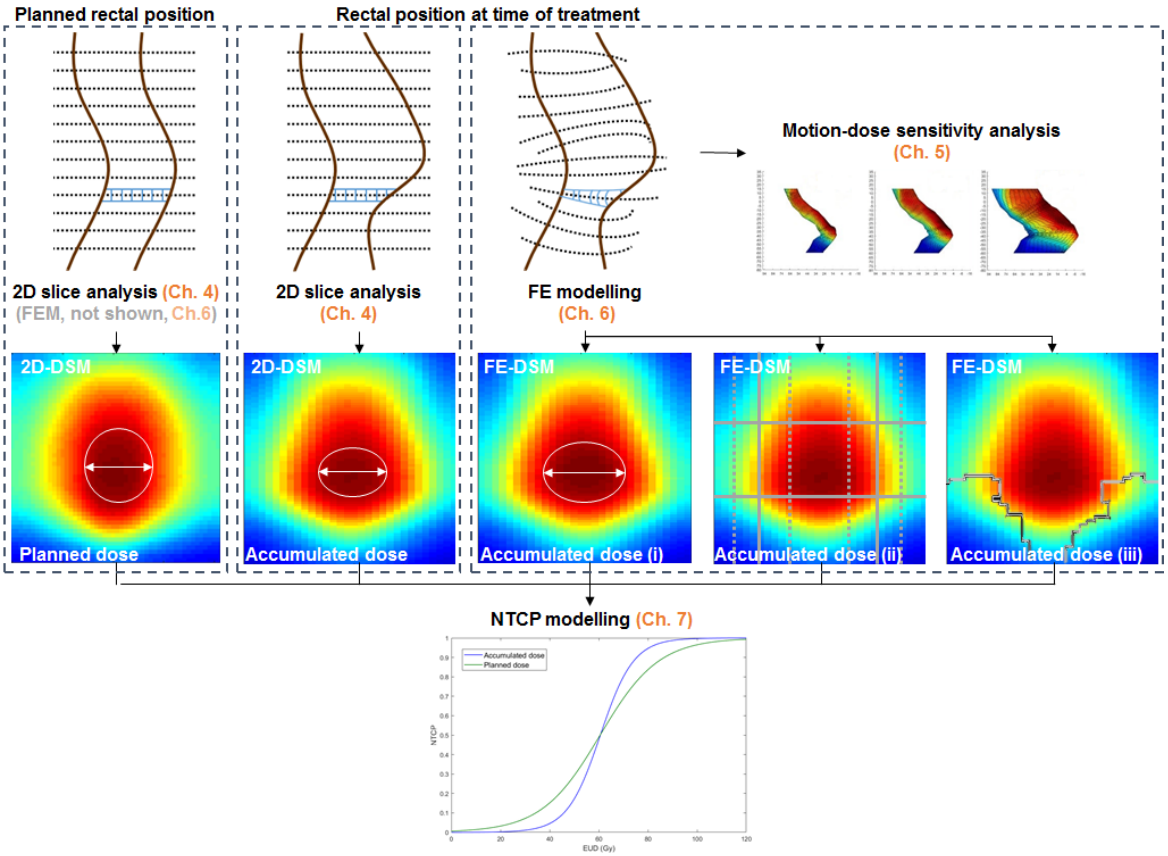


Fig. 1.3 Thesis overview; diagrammatic representation of the different analysis approaches presented in results chapters (Ch.). In Ch.4, 2D slicewise analysis of rectal contours is used to generate 2D dose surface maps (2D-DSMs) of the rectal wall, based on both planning CT and daily image guidance scans. 2D-DSMs of planned and accumulated delivered dose were parameterised using equivalent uniform dose (EUD) and DSM dose-widths, the lateral extent of an ellipse fitted to a given isodose level (white ellipse and arrow). Biomechanical finite element modelling (FEM) improves accuracy and resolution of dose calculation. Ch.5 presents a sensitivity study using FEM to simulate changes in rectal motion and quantify the dose impact. In Ch.6, finite element dose surface maps (FE-DSMs) of planned and accumulated dose are parameterised using (i) EUD and DSM dose-widths, (ii) geometric subregions (grey lines), (iii) probabilistic subregions (grey contour). In Ch.7, the outcomes of chapters 4 and 6 are input into multivariate normal tissue complication (NTCP) models, and validated. Discriminative ability is compared between planned and accumulated dose models with toxicity.

1.5 Overview of chapters

Chapter 1: Introduction to radiotherapy, treatment planning, the VoxTox research programme, and research aims;

Chapter 2: Literature review and critical analysis detailing the current scientific environment regarding rectal toxicity in prostate cancer radiotherapy;

Chapter 3: Description of the systems providing the foundations of VoxTox, multidisciplinary contributions, and the data handling involved in this research;

Chapter 4: Analysis of planned and accumulated dose to rectal wall using conventional 2D dose surface maps (DSMs) and investigating correlations with rectal toxicity;

Chapter 5: Applying biomechanical FE analysis to calculating dose to the rectal wall. Exploring the range of patient motion, and resulting sensitivity of dose and NTCP;

Chapter 6: Parameterising FE-DSMs for investigating voxel-level dose-toxicity associations;

Chapter 7: Developing multivariate NTCP models for predicting rectal toxicity, optimised for both planned and accumulated delivered dose;

Chapter 8: Conclusions of the research conducted within this thesis, and identifying areas of future work.

Chapter 2

Linking dose with toxicity in prostate radiotherapy

This chapter presents a systematic review of the relevant literature within the current research setting of rectal dose-toxicity associations in prostate radiotherapy. This is in the form of a summary and critical appraisal of several national and international clinical trials and studies which have examined dose fractionation effects in prostate radiotherapy. Dose to rectum is discussed in terms of spatial dose quantifiers beyond the dose volume histogram (DVH) approach. Voxel-based characterisation of rectal dose is introduced, alongside applications to different clinical sites and different methodologies. Finally, through this critical analysis, gaps in the literature are identified and translated into the research questions addressed in this thesis.

2.1 Clinical studies in prostate radiotherapy

Radiotherapy offers a highly effective treatment for prostate cancer. The 10 year outcome data of the Prostate Testing for Cancer and Treatment (ProtecT) trial [72] presented evidence from 1643 men (545 radiotherapy, 553 surgery, 545 active surveillance) showing equivalent rates of disease-free survival between radical prostatectomy and radiotherapy, with fewer adverse effects when treated with radiotherapy. Radiotherapy is a non-invasive treatment option so may be preferable to surgical intervention, particularly for patients with pre-existing contraindications to surgery. Higher rates of disease progression were observed for patients undergoing active surveillance.

2.1.1 Conventional treatment

Several clinical trials have investigated the effect of dose fractionation in prostate radiotherapy in terms of biochemical disease control and gastrointestinal toxicity. Studies comparing conventional dose fractionation schedules (dose per fraction between 1.8 - 2 Gy) have shown that dose escalation leads to improved biochemical progression-free survival (bPFS) [48, 55, 57, 98, 124, 130, 166]. One example conducted in the UK was the Medical Research Council (MRC) RT01 trial. Data were accrued for 843 men with stage T1b-T3a prostate cancer between 1998 and 2002. Patients were randomised to a standard-dose conventional plan (64 Gy in 32 fractions) or higher-dose conformal plan (74 Gy in 37 fractions) with radical intent [148]. The trial reported evidence that dose escalation significantly improved bPFS [56], and subsequently dose escalation was recommended as the standard of care for localised prostate cancer in the UK in 2008 [48].

However, a consequence of improved efficacy in disease outcome was increased incidence of acute and late toxicities [55, 57, 149]. Additionally, despite reporting improvements in disease-free survival, dose escalation did not translate to an increase in overall survival [99]. RT01 results indicated that although within clinically acceptable levels, a higher incidence of bowel dysfunction was reported in the dose-escalated arm after 5 years. Rectal bleeding was the most commonly reported side-effect, followed by proctitis and diarrhoea, which tended to increase in severity up to 3 years [149]. Ten-year follow-up data supported the early results, and found that dose escalation improved disease-free survival. However, emphasis was placed on balancing this benefit against the heightened risk of prolonged toxicity effects post-treatment.

2.1.2 IMRT and hypofractionation

A prospective study conducted by the Radiation Therapy Oncology Group (RTOG) compared three dimensional conformal radiotherapy with intensity modulated radiotherapy (IMRT) for prostate cancer. IMRT achieved significantly lower doses to the rectum and bladder, and consequently reported lower rates of late rectal toxicity [99]. Findings were consistent with the QUANTEC review [94], that large volumes of the rectum irradiated to high dose levels were associated with toxicity.

There have been several large-scale randomised trials conducted for hypofractionation (2.5 Gy per fraction and above) [10, 13, 56, 88, 129, 133]. The UK CHHiP trial (Conventional or Hypofractionated High-dose Intensity Modulated Radiotherapy in Prostate Cancer) [52, 54] is a non-inferiority trial that recruited 3216 prostate cancer patients from multiple centres between 2002 and 2011. Patients were randomised into one of three different fractionation

schedules: 74 Gy in 37 fractions (2 Gy per fraction, the existing standard of care in the UK at the time), 60 Gy in 20 fractions (3 Gy per fraction), or 57 Gy in 19 fractions (3 Gy per fraction). Evidence presented at 2-year [54] and 5-year [52] follow-up indicated that hypofractionated radiotherapy to 60 Gy was non-inferior to the 74 Gy arm. Isoeffectiveness was achieved in terms of bPFS with equivalent levels of cumulative toxicity incidence, but no significant difference was observed for overall survival. Non-inferiority was also observed in the Dutch HYPRO (Hypofractionated versus conventionally fractionated radiotherapy for patients with localised prostate cancer) study between 64.6 Gy in 19 fractions and 78 Gy in 39 fractions [10]; PROFIT (Prostate Fractionated Irradiation Trial) between 60 Gy in 20 fractions and 78 Gy in 39 fractions [37]; RTOG 0126 between 70.2 Gy in 39 fractions and 79.2 Gy in 44 fractions; and recent outcomes of the Scandinavian-led HYPO-RT-HC (hypofractionated radiotherapy of intermediate risk localised prostate cancer) trial have now shown non-inferiority of ultra-hypofractionated radiotherapy (42.7 Gy in 7 fractions compared with 78 Gy in 39 fractions) [167]. In 2016, as a consequence of the CHHiP recommendations and the evidence base for non-inferiority of hypofractionation, the standard of care for external-beam radiotherapy of localised prostate cancer was changed to 60 Gy in 20 fractions [108].

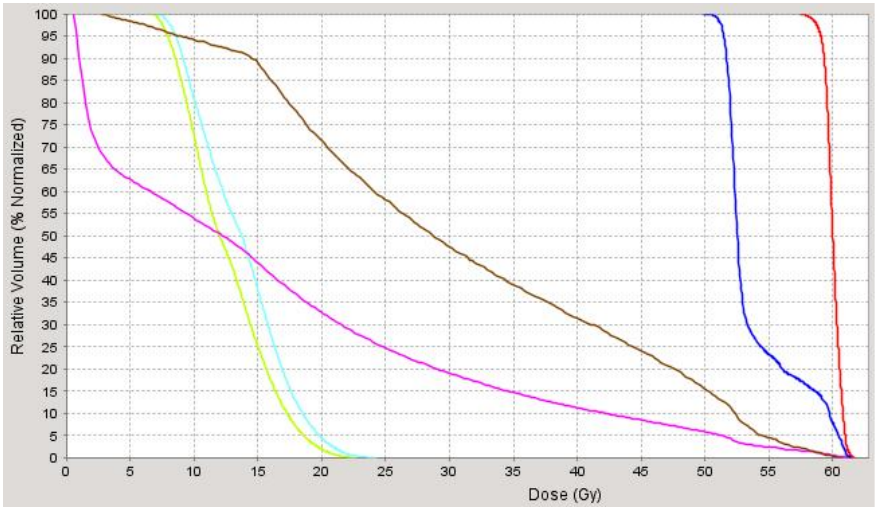
2.2 Quantifying dose to rectum

2.2.1 Dose volume histograms

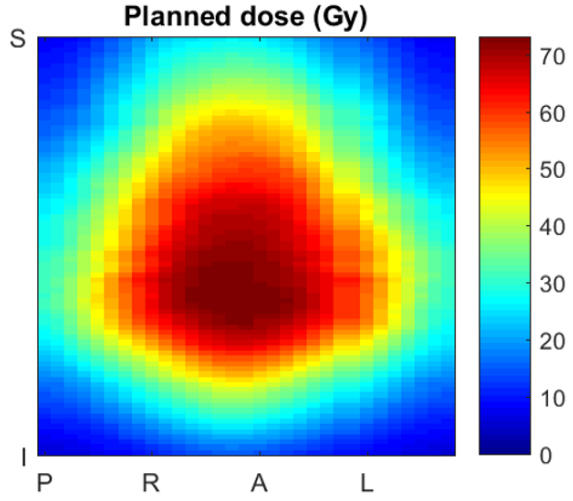
Dose volume histograms represent the cumulative frequency dose distribution to an organ (Figure 2.1a). DVHs are a powerful tool used in radiotherapy treatment planning to determine a clinically optimal treatment plan. The plan is iteratively optimised until a balance is achieved between coverage of the planning target volume (PTV) and organ at risk (OAR) sparing. This is evaluated by assessing whether DVH dose constraints have been met and compromising where necessary. However, existing DVH approaches have been criticised for lacking in information on spatial dose distribution [3, 171], and with high dose-gradient treatments such as IMRT, the geometric location of the high dose region becomes an increasingly important concern. Despite this limitation, they are widely used even though dose-volume constraints from 30 - 70 Gy have been found to be significantly associated with rectal toxicity [70], and high-doses in particular have been linked with rectal bleeding [87].

An additional clinical consideration not currently accounted for within the clinical trials discussed above is intra-organ radiosensitivity; all organs are assumed to be uniformly sensitive. For a serial organ such as the rectum, different regions may be more sensitive, and

some more tolerant, to the effects of radiation [4]. It has been demonstrated that discrete rectal toxicity endpoints have different underlying pathophysiologies, which may be attributed to different areas of the rectum. Data from the CHHiP trial have recently been interrogated to derive dose constraints optimised for each of four rectal toxicity endpoints [168]. The constraints were tighter than those applied in the CHHiP trial, which may result in reduced incidence of toxicity if applied in the future. However, testing has not yet been performed on the hypofractionated cohort.



(a)



(b)

Fig. 2.1 (a) Example Dose Volume Histogram (DVH); (b) Example Dose Surface Map (DSM)

2.2.2 Dose surface maps

Using data from the RT01 trial, approaches beyond the traditional DVH were investigated in order to characterise the shape and location of the dose distribution, and link with toxicity. Buettner et al. proposed the dose-surface map (DSM) as a representation of planned dose to the rectal wall [27]. By considering the rectal wall as a tubular surface that can be virtually cut and unwrapped, the dose could be visualised on a 2D plane. Generation of the DSM involved cutting the rectal contour at the posterior-most location and normalising longitudinally and circumferentially to a 21×21 matrix where each pixel represents the interpolated dose level. Visually, this creates a map with the high dose region positioned centrally, representing the anterior portion of the rectum closest to the prostate, with dose fall-off towards the edges of the DSM, as shown in Figure 2.1b.

Presenting dose to the rectal wall using DSMs allows characterisation of the shape and location of the dose distribution. Geometric features in the DSMs were extracted by computationally generating binary images corresponding to isodose clusters of discrete dose levels, and fitting an ellipse to the largest cluster. From the ellipse, measures of lateral extent, longitudinal extent, and eccentricity were recorded. Cumulative dose surface histograms (DSHs) were also generated for each DSM. Parameters were investigated for associations with rectal toxicity.

Results demonstrated several significant correlations between the proposed shape-based dose distribution features and various toxicity endpoints; in particular the lateral extent of medium to high isodose levels were significantly associated with rectal bleeding. The concluding recommendation of this study was that spatial dose patterns should be considered at the treatment planning stage

A follow-up study used sigmoidal functions fitted to histograms of dose-lateral-extent, dose-longitudinal-extent and dose-eccentricity, and used these as input for a machine learning algorithm [26]. A support vector machine was trained to predict toxicity based on the sigmoidal functions with the strongest predictive power for each clinical endpoint, i.e. rectal bleeding based on lateral extent, and so on. This allowed calculation of individual complication probabilities rather than binary outcomes, and patients could be ranked according to complication probability. Patterns were identified that correlated to low risk of complication, which improved upon DVH based models alone. Results verified the previous findings, that the lateral dose extent was found to be a significant predictor of toxicity.

The DSM approach to representing dose to the rectal wall has been applied in several studies interrogating the links between dose and toxicity [115, 139, 171]. Buettner's findings have subsequently been supported by independent studies that found rectal bleeding to be associated with the lateral extent of the mid to high isodoses [103, 115]. Moulton et al. [103]

also reported an association between the high doses and diarrhoea, stool frequency, and urgency, as well as further toxicity associations with spatial features of DSMs. Casares-Magaz et al. [34] found defecation urgency and faecal leakage to be associated with intermediate to high dose in the central to upper rectal wall. Wortel et al. [171] found links between several spatial DSM patterns with acute rectal toxicity.

By comparing the DSMs of patients with and without toxicity, it is possible to observe patterns and test for significance. Onjukka et al. [115] found that dose to the caudal extent of the treated volume and anal canal were associated with faecal incontinence. The authors recommended the inclusion of the anal canal as an OAR in treatment planning. Moulton et al. [103] suggest that reducing dose to the superior and inferior rectum may be important considerations for rectal pain and faecal urgency, respectively.

DSMs of the rectal wall have also been used to combine external beam radiotherapy doses with brachytherapy in the treatment of cervical cancer [42]. The DSM approach has also been applied to provide spatial dose information for the duodenum [169], and to the bladder to investigate spatial effects associated with bladder toxicity [121, 122].

Some considerations and limitations must be identified for the use of DSMs, Due to the relatively new adoption of the technique, few large scale studies have been conducted. There are no consensus guidelines on determining splitting point and normalisation for DSM generation; users have developed their own in-house systems. To date, DSM feature analysis has primarily focused on properties of ellipses fitted to isodose levels. However, alternative methods to characterising DSM dose distributions may be more beneficial for determining dose-toxicity associations. Further complexities arise when considering DSMs as a tool for accumulating dose, as rows from planned DSMs correspond to axial slices, and the tracking the same anatomical region for the same patient throughout the course of treatment may require methods for tracking extraplanar motion.

2.2.3 Dose line histograms

An alternative method has been used to parameterise rectal dose, the Dose Line Histogram (DLH). At each slice, a histogram of the dose to the rectal circumference is produced. Longitudinally, the histogram corresponds to the maximum length of each isodose level [73]. Using this approach on planning CT data only, strong DLH correlations were established. The authors acknowledge that improvements could be made to increase reliability of correlation models by accounting for interfraction motion.

2.3 Voxel-based analysis

For effective registration and modelling of intra and inter patient dose distributions, greater spatial accuracy is required than previously achievable using DVH or DSM approaches. Voxel-based analysis is a research focus that has been gaining momentum with the paradigm shift towards spatial dose analysis. Some advantages of voxel-based approaches are that they are not limited by imaging slice width, and do not require normalisation around the rectal circumference. Furthermore, voxelwise analysis offers a solution for extraplanar expansion and tracking which is closer to anatomical behaviour, allowing increased resolution of data, which in turn may reveal information on intra-organ inhomogeneous radiosensitivity.

2.3.1 Image-intensity based voxel mapping

A dose-image-based approach to inter-patient mapping was presented by Acosta et al [3, 116]. For 105 patients, the 3D dose distribution was mapped to a common template using a non-rigid hybrid organ/intensity registration method [41, 151]. A subregion was identified in a region of the anterior rectal wall, located within 15 mm of the prostate, where the absolute dose difference between patients with and without rectal bleeding was measured to be 6 Gy. The most significant mean-DVH differences were observed at the 40-65 Gy intermediate dose levels, corresponding with Buettner's [27] DSM findings.

A subsequent publication by the same group [61] exploits the voxel-level resolution of the model by dividing the rectum into rectal subregions at risk (SRR). Planned dose data at the voxel level were registered to a common template for 118 test patients. The rectal wall was divided into 'geometric', 'personalised', and 'generic' SRRs to explore spatial dose-toxicity associations. Subregions and toxicity correlations were investigated using a training set of 118 patients, and tested on a validation cohort of 53 patients. Validated results show that the dose to the inferior-anterior hemi-anorectum was highly predictive of rectal bleeding.

The voxel-based framework described above was used to develop toxicity prediction models for rectal bleeding using independent component analysis [64]. The model was generated using planned DVH parameters from training data, and has been successfully validated in a test cohort.

Similar voxel-based approaches have been used to determine spatial dose patterns associated with radiation-induced pneumonitis and fibrosis in lung cancer radiotherapy [118–120], and with urinary toxicity in prostate cancer radiotherapy [105]. The latter found dose to the urethra, and posterior and superior regions of the bladder, to be significantly associated with

different toxicity endpoints. The results for both urinary toxicity and rectal toxicity support the recommendation that spatial dose pattern should be considered at treatment planning.

2.3.2 Finite element modelling

Finite element modelling (FEM) can be used to predict anatomical organ motion and deformation by simulating biomechanical properties and tissue interactions. Organs can be deformably registered between two different image sets either on an individual organ basis according to material properties, or by surface interactions for multi-organ alignment [20]. These computational biomechanical models can account for the inhomogeneities and anisotropy associated with soft tissues undergoing large deformations [21], which cannot be accounted for by the approaches previously described.

The internal anatomy is discretised into individual deformable bodies with boundary conditions defined from the geometries of the clinical contours. Organs are assigned specific material properties, geometries, surface interactions, and the computer simulation is used to characterise the motion. FEM accuracy is more dependent upon the applied loading and boundary conditions than material properties [33], which is an important feature because material properties can be difficult to physically measure directly [21]. Additionally, unlike the solution presented by Acosta et al [3, 116] described in Section 2.3.1, FEM does not depend on image-intensity information which is known to be problematic in the case of the rectum due to the varying internal contents [143].

FEM can successfully model the intricacies of organ motion in 3D. By tracking of voxels according to biomechanical properties, we make the assumption that voxels are surrogates for groups of cells, and that the same cells remain in a particular voxel throughout the period of monitoring the organ deformation. However, the FE model cannot account for biological effects at the cellular level, e.g. cell cycle repair, repopulation, regeneration, or cell death. Furthermore, inter-voxel cell migration may be plausible, possibly in a similar phenomenon to the radiation 'bystander effect' whereby radiation-induced damage can be detected in unirradiated healthy cells away from primary the target site. Tracking of biological changes within deformable organs is difficult to measure, but could perhaps be better quantified and included into models in the future.

Biomechanical FEM provides a robust and anatomically-representative solution for 3D modelling and tracking of intra-organ variations [21]. Several studies have acknowledged the potential for significant improvements in radiotherapy treatment delivery by integrating biomechanical modelling into dosimetry tracking systems for adaptive radiotherapy.

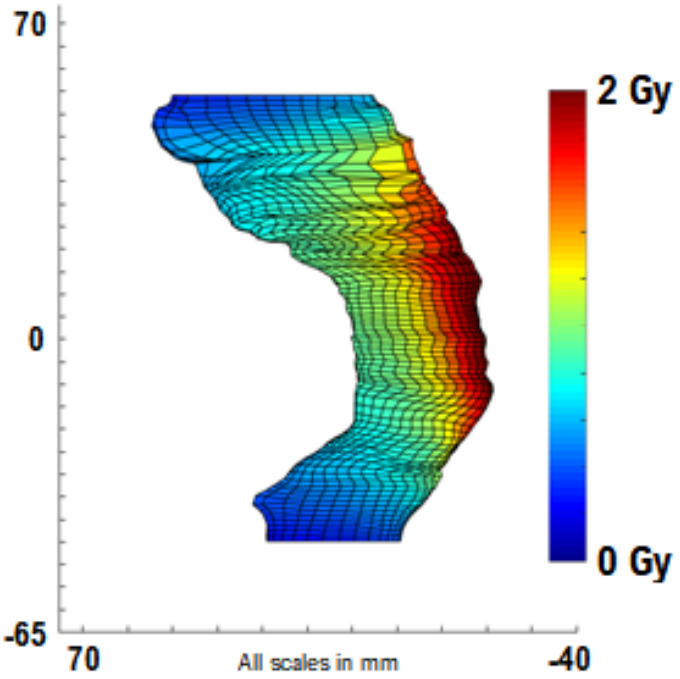


Fig. 2.2 Example of a biomechanical finite element model of the rectum based on the kVCT planning contour and dose per fraction.

2.3.2.1 Biomechanical FEM of the pelvis

Biomechanical FE modelling of the pelvic region has previously focused on studies of the prostate. Physically, the prostate is located immediately posterior to the bladder and anterior to the rectum, therefore its absolute position is not fixed within the body and can vary depending upon distension/filling/voiding of these adjacent organs [93]. In terms of prostate radiotherapy, not only does this motion pose a problem for image registration between scans acquired on different days and/or using different imaging modalities, but interfraction motion between fractions of radiotherapy treatment can introduce significant deviations between planned and delivered dose to the prostate and OARs. Image guided radiotherapy is used to ensure that the tumour is being accurately targeted at each fraction, and a key hypothesis of the VoxTox research programme is that these image guidance scans can be registered using FEM to accumulate delivered dose to the rectum.

Brock et al. [24] investigated the accuracy, sensitivity, and efficiency of contour variation and model parameters for FEM registration of the prostate. The accuracy was quantified by comparing residual error in the centre of mass of the implanted fiducial markers in the prostate, and the alignment of surface points, before and after deformation. Accuracy was found to be less than 1.5 mm in three-dimensions; this was less than the dimensions of the

image voxels, and therefore within acceptable limits for detectable accuracy. Sensitivity was determined by examining the change in accuracy for different scenarios of forward and backwards deformation, and was found not to be significant for different model resolutions or intraobserver variation in the anterior-posterior or right-left directions, with marginal differences detected in the superior-inferior direction (likely due to the effect of slice thickness in this direction). An increase in model size (number of elements used for surface mesh) by a factor of 24 was found to significantly increase calculation times from under a minute to over 30 minutes, without a significant improvement in model accuracy. This indicates that the level of accuracy associated with a higher resolution model, can in practice be achieved using a more efficient model that can be run in real-time using a reduced number of elements in the FEM calculation. The study was designed to provide evidence that FEM can be used to integrate multimodality imaging into radiotherapy treatment planning for accurate tumour definition. The conclusions are also applicable to registration of OARs such as the rectum.

An example of FEM being used to address the issue of organ mobility in the pelvic region was presented by Alterovitz et al. [9]. FEM was used to deformably register the prostate between two magnetic resonance imaging (MRI) scans acquired at different times; one acquired with an endorectal probe in-situ for diagnostic purposes, and the other without the probe for the purposes of radiotherapy treatment planning. This allows the specific disease site, as confirmed on a diagnostic quality scan, to be translated onto the scan of the patient in the actual treatment position, so that the area can be appropriately targeted. This may also be useful when considering the effect on prostate positioning when implanting rectal spacers or gels [62, 112, 128, 147, 159], which is a procedure that has recently been introduced clinically in the UK [109].

Boubaker et al. [18, 19] presented a predictive model for pelvic organ motion where the deformation of the prostate is dependent upon displacements of the neighbouring bladder and rectum, and is limited by fixed bony anatomy such as the sacrum and pubic symphysis. Extensive detail is provided for material properties, element types, thickness parameters, boundary conditions, and mesh generations. Urodynamic data were generated from experimental measurements from human cadavers [83], as well as porcine substrate [19]. As with the 2D-DSM approach in Section 2.2.2, a hollow-cylinder approach is adopted for the rectum. The rectal FE model is constructed of 3D shell elements of thickness 3 mm, whereas the prostate is modelled using tetrahedral solid-volume elements. Frictionless contact was assumed between rectum and prostate to simulate possible sliding between the organs. Rectal motion was up to 3 mm in the craniocaudal direction, and 5 - 20 mm in the anterior-posterior direction, as measured by Marchal et al. [93], and was fixed in place at the superior and inferior ends. A first order Ogden hyperelastic strain-energy model was used for the rectum,

assuming tissue homogeneity and isotropy. The mean intra-rectal inflation pressure was 1 kPa, applied uniformly to the innerpart of the rectal wall. Global error in rectal volume calculated at 12 displacement positions was 6 %. In the context of radiotherapy, this magnitude of change could make a large difference in the dose calculation, and hence estimation of toxicity risk. The author acknowledges that experimental discrepancies may exist between urodynamic properties in a living human, compared with ex-vivo measurements on a cadaver. However, the response of pig rectum tissue fell within the expected range of inter-subject variability [19]. Further experimental validation measurements to build a database for robust identification of constitutive laws were recommended. These would help to improve the predictive numerical FE platform so that the system could be considered a useful tool for adaptive radiotherapy.

Biomechanical properties of the rectum have been investigated using MRI imaging [36]. Internal rectal pressure was monitored using a pressure transducer with a balloon inflated in 50 ml increments. MRI images were acquired in order to quantify rectal wall thickness and elasticity. The maximum tolerable volume was 150 - 200 ml, corresponding to rectal pressure of 1.3 - 4.0 kPa. The rectal wall thickness was shown to vary across the rectum (1.8 - 4.0 mm); with increasing thickness from upper to lower rectum, and from anterior to posterior rectum. Strain also varied across the rectal wall, from 1.5 - 7.0 kPa from different areas. Biomechanical rectal properties varied on an intra- and inter-patient basis, which could perhaps reflect underlying radiation-induced toxicity issues. The author suggests that the results provide motivation for the inclusion of spatial properties into predictive response modelling for GI toxicity. This has been supported by other studies in the literature but further analysis may be required on larger patient cohorts due to relatively low incidence rates. Furthermore, the author proposes the inclusion of biomechanical properties into spatial evaluations of dose surface maps may better explain associations between particular subregions of the rectal wall and toxicity due to improved tracking capabilities offering heightened accuracy.

These studies into the biomechanical and material properties of the rectum have provided useful guidance for implementing FEM into the VoxTox workflow. Intra-rectal pressure has been shown to be difficult to measure and validate experimentally, which is why displacement-loading is used for many FEM anatomical simulations. It has been identified that biomechanical FEM can be used to identify regions of the rectum associated with rectal toxicity, with greater accuracy and resolution than existing approaches currently applied in clinical practice.

2.3.2.2 The Morfeus system

One of the most prolific systems for biomechanical finite element anatomical modelling in the current literature is an independent biomechanical-model based deformable image registration algorithm for multi-modality imaging, tissue tracking and dose calculation, known as Morfeus [20]. The Morfeus system has gone through many iterations and feasibility studies to improve accuracy and efficiency since its development in 2005 based on organ surface deformation in the Abaqus Computer Aided Engineering (CAE) Finite Element Analysis (FEA) environment (Dassault Systèmes) [20, 22]. The system has now been incorporated into commercial software [165] for abdomen, thorax, and pelvis.

Developed within the context of radiotherapy, for a solid structure example such as the liver, the initial anatomical contour is defined on the planning CT and used to generate an initial volumetric mesh to which material properties are applied. From a given daily kV cone-beam CT image guidance scan, a surface mesh is generated based on the liver contour of the day. Boundary conditions are then updated based on the new contour information, and the FE analysis is performed to determine a deformation map solution where the original volume is deformed to the new 3D volume. Through FEM, organ motion and interactions can be predicted, and the magnitude of the deformation can be used to quantify registration parameters between a reference and test image.

Deformation of the prostate has been successfully predicted in Morfeus using surface representations of the bladder and rectum to an accuracy of 2 mm, with the positioning of fiducial markers agreeing to less than the voxel resolution [76]. Sensitivity to mesh size was investigated for the prostate in order to minimise calculation time; where the calculations could be reduced to less than 1 minute, without significant impact on accuracy [24]. It was possible to reduce calculation time further, but at the expense of registration accuracy. Although sufficiently quick for on-treatment application in radiotherapy, in practice the system required anatomical contours as an input, and therefore the implementation of an autosegmentation system for image guidance scans was a vital development prior to implementation in order to be feasible within a clinical workflow [24].

Developing a model for the thoracic region involved many iterative cycles of complex analysis: methodical evaluation of element type, material properties and geometrical analysis [6]; friction and compressibility investigations [7, 5]; incorporating an image intensity-based hybrid system [135]; increasing the spatial resolution of the lung by applying boundary conditions to bronchioles [38]; and even reconstruction of 3D models from 2D planar images [110, 111]. The author suggests that the latter could be used to calculate historical treatments retrospectively to investigate late toxicity effects.

Morfeus has been applied for deformable image registration and FEM to retrospectively accumulate delivered dose for liver stereotactic body radiotherapy. A study explored the difference between planned and accumulated dose accounting for positioning, breathing and deformation effects which were found to range from -15 % to 5 % for tumours, and -42 % to 8 % for normal tissues [163]. This result prompted further investigation into the possibility of dose escalation to the tumour [164]. Margins surrounding the tumour were assessed using FEM and consequently, a smaller 'dose-probability-based PTV' was defined enabling dose escalation of 4 Gy to the tumour. Breathing motion was significantly different to predicted breathing for over 50 % of patients, and for the majority of patients, accumulated dose deviations were greater than 5% with respect to planned dose. These results were important because lower doses to the tumour in liver SBRT are associated with poorer survival and disease control. The authors suggest that accumulated delivered dose can be an important tool for adaptive radiotherapy, and may facilitate dose escalation to improve clinical outcomes in the future.

Dose accumulation as modelled using Morfeus has been verified experimentally using CT-based matching of a novel 3D deformable gel dosimeter [114]. Early models were based on MRI-contoured images, offering superior image quality over the CT and MVCT or CBCT image guidance scans [20]. However, validation results were within sub-millimetre agreement. The algorithm utilised external boundaries to perform the deformation registration, which was then confirmed by comparing planned and delivered dose distributions [162]. Though the experimental results seem promising, clinical validation was reportedly challenging due to difficulty in feature identification in low-contrast images.

The Morfeus system was translated from a development research platform into the routine online image guidance workflow at the developing centres. The system has recently been integrated into the Raysearch treatment planning system [165]. The system allows quantitative assessments of interfraction variability across many anatomical sites, demonstrating the versatility of FEM application in medical imaging. Accumulated dose calculations have been reported, however, the focus has been mainly on localisation and adaptive management. As yet, there have been no publications in the literature using FEM to link delivered dose to the rectum with toxicity in prostate radiotherapy.

2.4 Overcoming current limitations by accounting for interfraction motion of the rectum

A limitation of each of the studies and clinical trials discussed in this chapter so far is that the quoted dose to rectum is based upon the position at a single snapshot in time, the planning CT scan. This image of the internal anatomy is usually acquired several weeks prior to the beginning of the radiotherapy treatment, and although daily treatment setup images are qualitatively compared against the planning CT as a baseline, no accounting has been made for interfraction movement. However, rectal motion has been shown to be significant throughout the course of radiotherapy [117], and has the largest influence on the shape and motion of the prostate [19]. Further studies [79, 176] have attempted to approximate delivered dose from planned dose by applying a Gaussian filter to effectively blur the static dose to account for intrafraction and interfraction motion, cell migration, and the bystander effect. However, it is acknowledged that variations in daily rectal filling are random and could significantly impact the actual delivered dose, which cannot be accounted for by the blurring filter. Even with the use of image guidance in prostate radiotherapy, the position of the rectum has been shown to vary more than expected [136], which indicates that the dose received by the rectum in prostate radiotherapy is different to that planned.

Studies are beginning to emerge where the effect of interfraction motion is accounted for by calculating delivered dose. Thor et al. [155] calculated 'motion-inclusive' delivered dose for 38 patients who had 9 repeat CT scans acquired during treatment. Repeat CT scans were manually contoured by radiation oncologists and rigidly registered to the planning CT scan so that a new dose calculation could be performed. QUANTEC dose-volume constraints were investigated and compared [97] between planned and motion-inclusive dose, and toxicity associations were assessed using logistic regression. Rectal volumes were found to be smaller amongst patients with higher toxicity and the rectal volume receiving more than 67 Gy was significantly associated with GI morbidity using the motion-inclusive dose. The author acknowledges the potential benefit to quantifying delivered dose using spatial dose metrics such as the dose-surface map in order to be able to identify specific areas of the rectum associated with toxicity effects. Further studies into dose-toxicity associations using motion-inclusive delivered dose were recommended.

Scaife et al. [139] demonstrated using dose-surface histograms generated from DSMs that planned dose was different to delivered dose for all patients. Casares-Magaz et al. [35] performed estimates of motion-inclusive dose to rectum by using the average dose calculated using the cone-beam CT images from approximately 30% of fractions in a case-control study of 28 patients. Several studies acknowledge that the effects of interfraction motion of the

rectum have not been accounted for, and recommend research into this area as a worthwhile endeavour [25, 75, 96, 156, 173].

2.5 Predicting toxicity using accumulated dose

In 2010, the QUANTEC vision paper recommended the development of techniques for dose accumulation in order to improve the understanding of radiation-induced toxicity effects [94]. This landmark publication recognised that planned dose is not equal to delivered dose, and that with the increasingly high dose gradient techniques available, small deviations from the planned position can lead to larger differences in the dose actually received. Additionally, with the increasing use of higher dose-per-fraction delivery techniques such as stereotactic ablative body radiotherapy, daily anatomical deviations have a greater impact on the total delivered dose. Whereas previously the effects of interfraction motion would be averaged across 20 or 37 fractions, now a single fraction can represent a fifth of the total treatment delivery.

Research has been undertaken to improve NTCP modelling since the publication of this article, whereby spatial aspects or specific dose-volume constraints have been investigated in terms of principal component analysis [64] and machine learning neural networks [25, 177]. However, models based on delivered dose have yet to be developed. By relating delivered dose to toxicity, it may become possible to maximise the therapeutic ratio, advance adaptive radiotherapy, and improve treatment outcomes.

2.6 Identifying research questions

To summarise the scientific literature reviewed in this chapter, there have been extensive international studies exploring the effects of rectal toxicity in prostate radiotherapy. Toxicity remains a clinical issue in radiotherapy and the underlying mechanisms are fairly poorly understood. Existing toxicity prediction models seem insufficient for the modern high-dose gradient complex treatments currently being delivered, as the historic data used to derive constraints for dose-limiting organs were based on 3D conformal techniques. Nevertheless, these DVH constraints are commonly applied in radiotherapy treatment planning to achieve a compromise between tumour coverage and healthy organ sparing. As acknowledged by many authors, DVH parameters are limited by their lack of any spatial dose dimension. The search for more appropriate dose descriptors for the rectal wall seems to be focused on parameterisation of dose surface maps. An overarching limitation and recommendation of studies examining dose-toxicity associations in prostate radiotherapy, has been that motion-

inclusive dosimetric data is required for a more complete understanding. Radiation treatments to the pelvis region are particularly susceptible to interfraction motion which can lead to large deviations between planned dose and that actually delivered. This introduces an inherent uncertainty in dose-toxicity modelling based on planned dose.

Parallel advances have seen the application of the finite element approach to anatomical modelling, which facilitates 3D tracking of organ motion. Rather than image intensity-based deformable registration tools that can be limited to planar expansions or struggle at large deformations, FEM applies constitutive biomechanical laws to anatomical simulations. This provides the capability to track voxel-level extraplanar deformations, which can improve the accuracy and resolution within a system for the calculation of delivered dose. FE models have previously been applied for accumulating daily dose to the liver, but not yet to the rectum in prostate radiotherapy.

Here, we seek to combine gaps in the literature to explore dose-toxicity associations with delivered dose to rectum, explore the application of FEM to the rectum in terms of sensitivity, interrogate the voxel levels resolution of FE rectal simulations for patterns of spatial dose asscociated with toxicity, and finally, culminate this data with patient factors into multivariate normal tissue complication probability models.

This PhD research has been conducted as part of the VoxTox research programme, linking dose at the voxel level with toxicity. The hypothesis of the study is that delivered dose is a better predictor of toxicity than planned dose. The focus of the research conducted within this thesis is on delivered dose to the rectum as a predictor of rectal toxicity in prostate radiotherapy. The mechanisms for calculating delivered dose to a 2D-DSM were already in place prior to the commencement of this research. The areas of research that have been identified as missing from the literature, and are being addressed as original scientific contributions within this thesis, are:

- The use of spatial aspects of delivered dose to rectum, represented using 2D-DSMs, to investigate associations with toxicity
- Optimisation of a biomechanical FE model for the purposes of accumulating delivered dose to the rectum
- Application of FE modelling to a sensitivity study to improve knowledge of the relationship between rectal motion and dose
- Exploiting the voxelwise nature of FEM dose accumulation to interrogate dose-toxicity associations at a higher resolution using subregion analysis
- Development of NTCP models based on spatial aspects of delivered dose to rectum

Chapter 3

Data management in VoxTox

In this chapter, the intricacies of the VoxTox data and workflow are described. The following processes were developed by others prior to commencement of this PhD project: autosegmentation of the rectum on megavoltage (MV) image guided radiotherapy (IGRT) scans, independent dose calculation of image guidance scans, calculation of two-dimensional dose surface maps of the rectal wall, anonymisation of patient data, automated processing system, collection of baseline and toxicity data, and coding for a 3D finite element rectal simulation from experimental data. These systems provided the foundations for the work conducted within this thesis. Appropriate accreditation for the work undertaken in developing these systems is acknowledged within this chapter, and further summarised in Appendix A, Statement of originality.

Within this PhD project, further modification of several processing systems and manipulation of data were required to address the research question, 'Is delivered dose to the rectum a better predictor of toxicity than planned dose in prostate radiotherapy?'. These included: improvements and updates to the autosegmentation algorithm, parameterisation of 2D-DSMs for calculation of equivalent uniform dose and dose-widths, combining fractionation schedules to equivalent dose in 37 fractions, interpretation and processing of toxicity data, and translation of the experimental 3D finite element model to a model appropriate for handling patient DICOM data. Each of these stages involved collaborative working and have been accredited as such in the relevant sections.

3.1 VoxTox infrastructure

Radiotherapy is an advanced technology for cancer treatment involving large quantities of complex data; from imaging systems and planning algorithms, to treatment delivery and verification. The VoxTox dataset is unique in that it incorporates full treatment planning data

(CT images, structures, plan, dose) alongside daily treatment data (image guidance scans, structures, dose) and toxicity data (baseline, acute, late), all prospectively collected for 874 patients (Table 3.1). The VoxTox project was launched in 2013, and it took several years to develop systems capable of handling and processing such large volumes of patient imaging, structure, dose, and outcome data.

Table 3.1 VoxTox patient numbers for 'discovery' and 'consolidation' cohorts across 3 clinical sites (prostate, head and neck, and central nervous system, CNS), treated between 2013 and 2018, as recorded May 2018.

	Prostate	Head & Neck	CNS
Discovery	268	95	-
Consolidation	267	220	24
Total (n = 874)	535	315	24

The VoxTox dataset consists of two discrete patient cohorts, 'discovery' and 'consolidation'. The discovery cohort patients were patients who had previously been treated and were recruited retrospectively. Full imaging and dosimetric information were available. Toxicity data were collected prospectively, but baseline data were sparse and, where available, were extracted manually from clinical notes. Furthermore, these patients were treated prior to the implementation of formalised VoxTox contouring and imaging protocols. The consolidation cohort patients were recruited prospectively, at diagnosis. Detailed toxicity data, including baseline, acute, and late symptoms, were recorded using electronic case reporting forms. Standardised protocols were followed for contouring and imaging.

VoxTox patients were treated using TomoTherapy, a dedicated image-guidance intensity-modulated radiotherapy system. The Hi-ART TomoTherapy linear accelerator is a 6 MV linear accelerator on a rotating CT-type ring-gantry. The patient is positioned on one of two TomoTherapy Hi-ART units, on the treatment couch, and megavoltage image guidance scans are acquired to ensure accurate targeting of the tumour. Any required positional corrections are made based on the daily images. The radiation treatment plan is then delivered. Mechanical multi-leaf collimators shape the fan beam to deliver helical intensity modulated radiotherapy (IMRT).

3.1.1 Megavoltage image guidance scans

A Xenon-filled CT detector is mounted on the ring-gantry, directly opposing the linear accelerator, providing megavoltage computed tomography (MVCT) imaging capabilities. The beam energy is reduced to 3.5 MV to enhance soft-tissue contrast and reduce imaging dose to the patient to less than 3 cGy [174]. The MVCT field of view is shorter than the kVCT

planning scan field of view which limits the imaging dose, yet is sufficient for positional verification using soft tissue matching. The 'coarse' image setting is used, providing imaging slices every 6 mm. Compared to kVCT imaging, the MVCT image quality is poorer due to lower contrast at soft tissue boundaries and reduced signal to noise ratio.

For each treatment fraction, an MVCT image is acquired immediately prior to treatment delivery for the purposes of online target localisation. Patient positioning and anatomy is compared with the reference planning kVCT image to visually verify target accuracy. Small rigid translations of the treatment couch (anterior-posterior, superior-inferior, left-right) or gantry angle (roll) may be applied to register the MVCT with the kVCT and align the anatomy. Motion during treatment, or intrafraction motion, was not possible to monitor and is outwith the scope of this investigation. For prostate radiotherapy, image registration is prioritised for the planning target volume (PTV). IGRT scans are inspected for rectal dilation and if deemed excessive, remedial action is taken prior to treatment delivery. However, other than qualitative visual inspection of rectum and bladder volumes, variations in size, shape, or position of organs at risk (OARs) are generally not taken into consideration.

The VoxTox project identified that the full potential and value of information available within IGRT images was being unrealised. Important information on daily positioning of, and hence true dose being received by, OARs at time of treatment is generally disregarded in routine clinical practice. Within VoxTox, systems have been developed for automatic detection of OARs on MVCT scans, and independent calculation of dose based on the Hounsfield Units (HU) of the MVCT scan. This enables quantitative information to be extracted from MV IGRT scans, which facilitates the calculation of delivered dose to the OARs.

3.1.2 Calculating delivered dose

An independent dose calculation software, CheckTomo, was internally developed at Cambridge University Hospitals by Dr Simon Thomas [152, 154]. CheckTomo is a ray-tracing dose calculation algorithm, utilising the HU values directly from the MVCT images. The MVCT HU values have been shown to be sufficiently stable for accurate dose calculation, and imaging systems are calibrated weekly. Initially developed for the purposes of treatment plan verification, the system is capable of calculating full 3D volumetric dose distributions [154].

In the first iteration of dose-toxicity analysis, accumulated dose calculated using CheckTomo was compared with planned dose from the TomoTherapy treatment planning system. Further advances were made so that a comparison could be made between planned and accumulated dose, both calculated using CheckTomo. Comparative measurements were acquired at each stage to test the effect of updating the algorithms.

3.1.3 Autosegmentation of the rectum

The development of robust and automated approaches to segmentation has been identified as a key aspect in the pursuit of delivered dose calculation for adaptive radiotherapy [82], as manual contouring of daily image guidance scans is unfeasible for large volumes of patients. Autosegmentation was an essential tool for the VoxTox study; with over 500 prostate patients, each with 20 or 37 IGRT scans each, and an average of 12 slices per scan, it would have been unfeasible to manually contour over 14,000 image sets, totalling over 170,000 image slices.

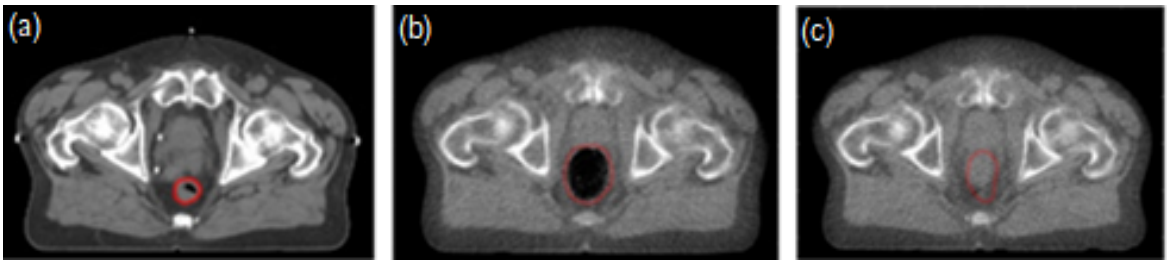


Fig. 3.1 Rectal contours (red); (a) manually contoured on the kVCT scan, (b) autosegmented on a daily MVCT image guidance scan, showing a larger contour with air cavities, (c) autosegmented contour on another daily MVCT image guidance scan, showing another change in contour with respect to baseline due to increased rectal contents.

A common approach in the literature for identifying contours from IGRT is through deformable image registration (DIR). The anatomy can be morphed from the image of the day to the template planning CT scan via deformation vectors. Registering anatomy to the planning scan can allow dose calculation to the organs using the treatment planning system. However, a problem when considering the application of DIR to the rectum is that it undergoes large and unpredictable spatial deformations caused by rectal contents and intestinal gas [97, 114, 136, 161]. Common DIR algorithms struggle due to intensity variations and the lack of one-to-one correspondence between a full, gassy, or empty rectum [40, 114, 175]. Previous studies investigating the dosimetric effects of interfraction rectal motion have been dependent upon manual delineation of the rectum on IG scans [11, 43, 44, 75, 85, 96, 123, 125, 145], consequently being limited in sample size. One approach attempting to address this limitation was to apply a blurring filter and implement statistical simulations to estimate motion-inclusive delivered dose [155]. A common recommendation of these studies was the development of robust systems for autosegmentation of the rectum, as a crucial component towards achieving automated adaptive radiotherapy for prostate patients.

The algorithm developed for autosegmentation of the rectum on MVCT scans within the VoxTox study was developed by Prof Michael Sutcliffe, Cambridge University Department of Engineering. The algorithm is a primary segmentation system based on a modified Chan-Vese algorithm [39]. Pre-processing steps involve HU/intensity scaling, identifying search

regions, dealing with air pockets, and handling the prostate. Optimal values were derived for smoothing, weighting, and bias parameters. For each slice, the starting point and search region are defined by the planning contour from the original kVCT. HU values in the soft tissue range corresponding to rectal material (-10 - 100 HU), as well as regions identified as air pockets ($<$ -130 HU), are rescaled and assigned a greyscale intensity value between 0 - 1. The algorithm then identifies the rectal contour in 2D. Post-processing in 3D is used to detect erroneous contours (often in regions of low contrast-to-noise such as the lower-third), and improves these either via interpolation or direct replacement with the original kVCT planning contour as a best estimate.

Within this PhD, work was done to identify areas of improvement in the autosegmentation algorithm. A small study was undertaken to identify erroneous contours, which occurred particularly around regions of air pockets and at large diameters. The results of this study were reported back to Prof Sutcliffe and improvements were implemented within the code. With each iteration, dosimetric evaluations were conducted to assess the effect, and fed back to VoxTox colleagues. The final version used for reporting all patient results for the consolidation cohort was version 1.6. The details of version 1.6 of the autosegmentation algorithm were published jointly by Shelley & Sutcliffe [143] in the journal of Biomedical Physics & Engineering Express.

3.1.4 Automated processing

Plan, imaging, contour, and dose data were stored as Digital Imaging and Communications in Medicine (DICOM) files and tokenised to VoxTox identifiable labels [132]. Down-sampled kVCT images were extracted from the TomoTherapy archive. Once removed of patient identifiable information, the data were transferred from the hospital to the University of Cambridge Department of Physics. Batch processing was performed using the computational task-management system GANGA (Gaudi and Grid Alliance) [102, 154]. Automated processing was an essential element of the VoxTox workflow, allowing large-scale calculation of patient data, independent of clinical systems. Total data volume for 870 patients amounted to approximately 250 GB and was stored on the Large Hadron Collider (LHC) Tier 2 Centre at the University of Cambridge Cavendish Laboratory.

3.1.5 Recording toxicity

With ever improving disease control [53, 50] and survival rates [32], post-treatment quality of life becomes an increasingly significant consideration during treatment planning, alongside target coverage. Despite technological advances allowing improved tumour targeting and

avoidance of organs at risk due to the tight dose conformity achieved through intensity and volumetric modulated, rectal toxicity in prostate radiotherapy remains a clinical issue, with rates reported in the literature of up to 22 % for late rectal bleeding greater than grade 2 [71, 87].

Interpretation of toxicity studies in the literature can be difficult due to the plethora of different reporting systems available [77]. For example, rectal toxicity outcomes have been reported using scoring systems such as Common Terminology Criteria for Adverse Events (CTCAE) [158], Late Effects of Normal Tissues/Subjective, Objective, Management, Analytic (LENT-SOMA) [1], University of California, Los Angeles, Prostate Cancer Index (UCLA-PCI) [89], Radiation Therapy Oncology Group (RTOG) [127], as well as outcome measures developed locally from reporting institutes. Furthermore, definitions of symptoms and grading intervals can differ between reporting systems. Toxicity is considered to be noisy data, as conditions and severity can be subjective to the patient and/or reporting clinical staff.

A toxicity questionnaire was purpose-designed for the VoxTox study by Dr Jessica Scaife and Amy Bates, Lead Research Radiographer, which encompassed all observational toxicity endpoints reported in prostate radiotherapy trials in UK within the previous 15 years [137]. Electronic case reporting forms (eCRF) and patient reported outcome measures (PROMs) were used for prospective data collection of baseline (prior to start of treatment), acute (during treatment and up to 90 days post-treatment), and late (beyond 90 days post-treatment) toxicity incidence. Toxicity was assessed via radiographer interview and patients were followed for 5 years following treatment.

For analysis of the consolidation cohort, 7 clinical endpoints were investigated, as shown in Table 3.2. The range of toxicity endpoints investigated were selected based on the most commonly reported toxicity outcomes presented in the literature [70, 77]. For each endpoint, increasing grades of severity were included where the incidence rate was greater than approximately 10%. In general, grade 1 toxicity refers to asymptomatic or mild symptoms with no intervention required, grade 2 indicates moderate symptoms requiring clinical intervention, and grade 3 denotes severe symptoms. Within the VoxTox dataset, no patients reported grade 4 toxicity or above.

Table 3.2 Definitions of toxicity endpoints investigated in the VoxTox consolidation cohort dataset. The grading system refers to increasing severity. CTCAE = Common Terminology Criteria for Adverse Events v4.03 [158], RTOG = Radiation Therapy Oncology Group [127], RMH = Royal Marsden Hospital [70], UCLA-PCI = University of California, Los Angeles, Prostate Cancer Index [89]. A semicolon indicates 'or', and a dash indicates that a grade is not available. ADL = activities of daily living. Grade 0 would indicate no reportable toxicity.

Toxicity endpoint (Scoring system)	Grade 1	Grade 2	Grade 3	Grade 4	Grade 5
Diarrhoea (CTCAE)	Increase of $<$4 stools per day; mild increase in ostomy output	Increase of 4 - 6 stools per day; moderate increase in ostomy output	Increase of $\geq$ 7 stools per day; incontinence; hospitalisation; severe increase in ostomy output; limiting self care ADL	Life-threatening consequences; urgent intervention indicated	Death
Faecal incontinence (CTCAE)	Occasional use of pads required	Occasional use of pads required	Severe symptoms; elective operative intervention indicated	–	–
Proctitis (CTCAE)	Rectal discomfort; intervention not indicated	Symptoms (e.g., rectal discomfort, passing blood or mucus); medical intervention indicated; limiting instrumental ADL	Severe symptoms; faecal urgency or stool incontinence; limiting self care ADL	Life-threatening consequences; urgent intervention indicated	Death
Rectal haemorrhage (CTCAE)	Mild; intervention not indicated	Moderate symptoms; medical intervention or minor cauterisation indicated	Transfusion, radiologic, endoscopic, or elective operative intervention indicated	Life-threatening consequences; urgent intervention indicated	Death
GI toxicity (RTOG)	Mild diarrhoea; mild cramping; bowel movement 5 times daily; slight rectal discharge or bleeding	Moderate diarrhoea and colic; bowel movement $>$5 times daily; excessive rectal mucus or intermittent bleeding	Obstruction or bleeding; requiring surgery	Necrosis or perforation; fistula	Death
Bowel frequency (RMH)	3-4 times in 24 hours	5-8 times in 24 hours	$>$8 times in 24 hours	Life-threatening consequences; urgent intervention indicated	Stoma in situ
Bowel bother (UCLA-PCI/ RMH)	Very small problem	Small problem	Moderate problem	Big problem	–

3.2 Interpreting rectal toxicity

The processing of patient reported toxicity data was not fully automated, and required manual intervention in order to be output in a useful format. Data from PROMs recorded within the eCRFs, were exported in a raw format. Translating PROMs into formalised clinical endpoints requires deep clinical understanding and interpretation. The mapping rules from raw questionnaire answers to populating standardised reporting systems were externally verified by Dr Agnieszka Lemanska, University of Surrey. The process of extracting raw toxicity data and translating into standardised reporting formats ready for analysis involved several stages and depended upon the input of Amy Bates, Lead Research Radiographer, Cambridge University Hospitals, Karl Harrison, Research Associate, Cavendish Laboratory, and Dr David Noble, Clinical Research Fellow, Cambridge University Hospitals. The final interpretation of the data and combining with dosimetric and baseline data was performed within the scope of this PhD.

The VoxTox research programme ceased patient recruitment in May 2018. The data analysed in this thesis for the consolidation cohort are from the patient spreadsheet frozen in January 2018. This included 186 patients with full dose information, baseline toxicity, and a minimum of 2 year follow up, with complete data entries at 6 months, 1 year, and 2 years. Of these, 110 patients were prescribed 74 Gy, and 76 patients prescribed 60 Gy.

The clinical endpoints reported for rectal toxicity within this thesis covered the most commonly reported toxicity endpoint in the literature and included: CTCAE Diarrhoea [158], CTCAE faecal incontinence [158], CTCAE proctitis [158], CTCAE rectal bleeding [158], UCLA-PCI bowel bother [89], LENT-SOMA stool frequency [1], and RTOG gastrointestinal (GI) toxicity [127]. For bowel bother, UCLA reporting system has been translated to the Royal Marsden Hospital (RMH) scale [70] so that increasing severity is reflected in an increase in numerical scale (from 0 to 5), similar to other toxicity endpoints being reported. Incidence rates are presented for all consolidation cohort patients in Figure 3.2. The toxicity rates recorded within the VoxTox trial were generally on the high side of the ranges reported in the literature. This could be due to the study commencing prior to CHHiP [57] dose constraints being applied in treatment planning, or to the heightened sensitivity of reporting facilitated by the design of the patient questionnaires. However, the rate of bowel toxicity $\geq$Grade 2 (CTCAE), 5 year cumulative incidence, amongst VoxTox prostate patients was 17%, which falls within the RTOG bowel toxicity $\geq$ Grade 2 (RTOG) range of 13.7–24.9% for IMRT over the same timeframe, reported by Dearnaley et al. [52] and Wortel et al. [171], respectively. These rates of incidence indicate that toxicity remains an important clinical issue.

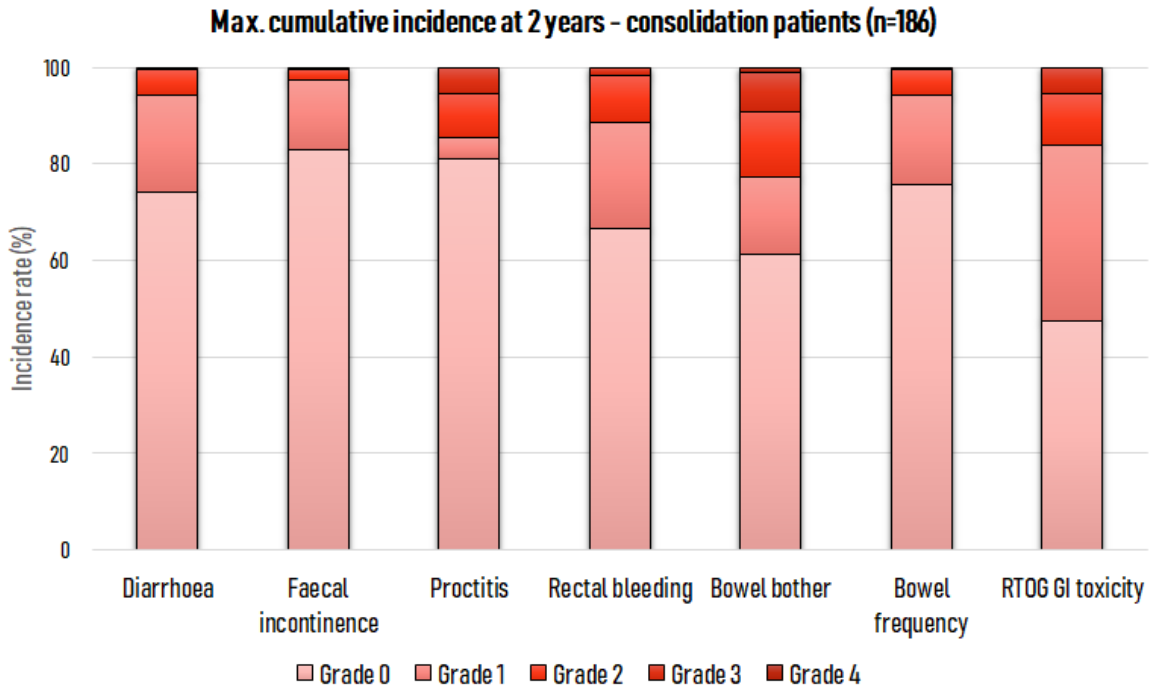


Fig. 3.2 Cumulative incidence of rectal toxicity at 2 years for consolidation cohort patients (March 2018). Grade 0 indicates no toxicity.

3.3 Combining fractionation schedules

The standard dose prescription for men with localised prostate cancer receiving radical radiotherapy in the UK was updated in 2016 as a result of evidence published by the CHHiP trial [52, 108]. The 5-year follow-up analysis demonstrated that 60 Gy delivered in 20 fractions of 3 Gy was non-inferior in terms of biochemical control, with non-significant differences in toxicity incidence, than the previous recommendation of 74 Gy delivered in 37 fractions of 2 Gy. Of 497 VoxTox prostate patients (as of 5th April 2017), 34% (n = 171) were prescribed 60 Gy in 20 fractions, and 66% (n = 326) were prescribed 74 Gy in 37 fractions (60/192 of 252 discovery and 111/134 of 245 consolidation patients, prescribed 60/74 Gy, respectively). Because the CHHiP trial found the two prescription regimes to be equieffective, VoxTox data were combined into a single dataset for toxicity analysis.

In the clinic, radiobiological equations are used to compare or combine different fractionation schedules. The biologically effective dose (BED), Equation 3.1, is a common measure which accounts for dose per fraction (d), number of fractions (n), and total dose (D), as well as the tissue-specific biological response, characterised by the α/β ratio. The constants α and β determine the shape of the cell survival curve in the linear-quadratic model, where

surviving fraction $SF = exp(\alpha D + \beta D^2)$. Typically, the α/β ratio is around 3 Gy for normal tissues and 10 Gy for solid tumours.

$$BED = nd\left(1 + \frac{d}{\alpha/\beta}\right) \tag{3.1}$$

The metric commonly quoted in clinical practice for combining different fractionation regimes is the Equivalent Dose in 2 Gy per fraction (EQD_{2Gy}),

$$EQD_{2Gy} = \frac{BED}{1 + \frac{2}{\alpha/\beta}} \tag{3.2}$$

However, for combining the VoxTox data it was not desirable to obtain the EQD_{2Gy}, but instead convert the 60 Gy fractionation to Equivalent Dose in 37 fractions, $EQD_{37\#}$. This is because although 60 Gy in 20 fractions is now the recommended dose prescription in the UK, within this thesis fractionations have been converted to 37 fraction equivalence in order to retain validity and applicability within the wider international context, where generally doses of up to 84 Gy are delivered over a longer treatment course duration. Furthermore, the prescription dose to the tumour is distinct from the inhomogeneous doses received by the rectum, whereas the fractionation schedule has a direct impact on the dose response. This approach was also followed by Holyoake et al. for modelling duodenum toxicity [78], and is more appropriate when combining and analysing equivalent effects.

Converting to $EQD_{37\#}$ involved solving Equation 3.1 for d, where n = 37, using the quadratic equation:

$$EQD_{37\#} = \frac{-1 + \sqrt{1 + \left(\frac{4 \times BED}{37 \times (\alpha/\beta)}\right)}}{\left(\frac{2}{37 \times (\alpha/\beta)}\right)} \tag{3.3}$$

The α/β ratio for late response of the rectum often reported in the literature is 3 Gy [97]. This value assumes a uniform dose distribution received by the whole organ which is not the case for modern IMRT so may not be appropriate for combining VoxTox dose data. Studies in the literature have quoted derived α/β ranges from 1.5 to 4.8 [58, 31, 95]. For VoxTox patient analysis, a study-specific α/β ratio was derived by fitting the data to CHHiP dose constraints and assuming equieffectivness between the two fractionation schedules [52]. The Equivalent Uniform Dose (EUD) for all dose bins up to 74 Gy was calculated to be 64.4 Gy, using Equation 3.4. The α/β ratio was iteratively refined until the EUD of the 60 Gy prescription converted to $EQD_{37\#}$ was also equal to 64.4 Gy. The optimal α/β ratio was found to be 2.149. This was close to the optimised α/β ratio of 2.3 Gy determined by

Table 3.3 Sensitivity of dose-toxicity association to change in α/β ratio. Comparison of EUD discriminative ability for CTCAE grade 2 rectal bleeding, quantified using area under the receiver operator curve (AUC) with corresponding 95% confidence intervals (CI). Presented α/β ratios representative sample within range reported in the literature.

α/β	AUC	95% CI
1.5	0.665	[0.550, 0.780]
2.149	0.638	[0.527, 0.749]
2.5	0.622	[0.509, 0.735]
3	0.613	[0.496, 0.729]
4	0.565	[0.440, 0.689]
4.8	0.547	[0.415, 0.680]

Marzi et al. [95]. Using an α/β ratio of 2.149, the $EQD_{37\#}$ of 60 Gy in 20 fractions was calculated as 74.31 Gy, and the final BED calculations for 60 Gy in 20 fractions and 74 Gy in 37 fractions were 143.76 Gy and 142.87 Gy, respectively, agreeing to within 1 Gy (<1%). By comparison, using the standard α/β ratio of 3 produced a difference of 1.5 Gy in EUD calculation between the two fractionation regimes, and a difference of 3.3 Gy (2.3%) between BEDs .

To examine the sensitivity of dose-toxicity association to a change in α/β, a representative sample of values within the range reported in the literature was used to calculate $EQD_{37\#}$. The effect on the magnitude of dose-toxicity association between EUD and CTCAE grade 2 rectal bleeding was compared using the area under the receiver operator characteristic curve (AUC). Results are shown in in Table 3.3. The extremes of the range of α/β ratios reported in the literature (1.5 v 4.8) can lead to a 12% discrepancy in reported AUC, where one could be considered predictive of rectal bleeding, and the other considered insignificant. However, the VoxTox derived value of 2.149 and the commonly accepted value of 3 yielded a difference in AUC of only 2.5%, where overlapping confidence intervals suggest that the difference between these results could be non-significant. Due to the consideration of the inhomogeneous dose distribution received by the rectal wall in determining the VoxTox-derived α/β ratio, the value of 2.149 is used for analysis of the combined cohort throughout this thesis. Once these values and equations had been determined the DSMs for patients prescribed 60 Gy in 20 fractions were corrected to $EQD_{37\#}$. Combining the datasets this way allowed a larger patient population to be analysed.

To generate $EQD_{37\#}$, EUD was equalised between 60 and 74 Gy fractionation regimes. EUD represents the uniform dose which would lead to equivalent probability of normal tissue injury as the corresponding inhomogeneous dose distribution [66]. As shown in Equation 3.4,

EUD is a form of weighted mean, calculated over N unit increments, dependent upon model parameter n, where $a = 1/n$.

$$EUD = \left(\frac{1}{N}\sum_i d_i^a\right)^{\frac{1}{a}} \tag{3.4}$$

For gastrointestinal toxicity, n has been commonly accepted as 0.09, but ranges from 0.03 to 0.5 have been applied in the literature [87]. To investigate an optimal value for n based on VoxTox data, the sensitivity of dose-toxicity correlation to parameter n was investigated for $0.02 \leq n \leq 1$, corresponding to $a \in \{50, 1\}$. Toxicity associations were assessed between EUD and CTCAE grade 2 rectal bleeding, as well as CTCAE grade 2 proctitis, and corresponding AUC values are presented in Figure 3.3. AUC results were found to plateau for both rectal bleeding and proctitis at approximately a = 11, corresponding to n = 0.09, before a slow AUC decrease at $a > 25$. Above a = 5, AUC varied by less than approximately $\pm$ 1 %. Therefore, in accordance with previous findings in the literature, the value a = 11.11(n = 0.09) was retained for all EUD calculations throughout this thesis.

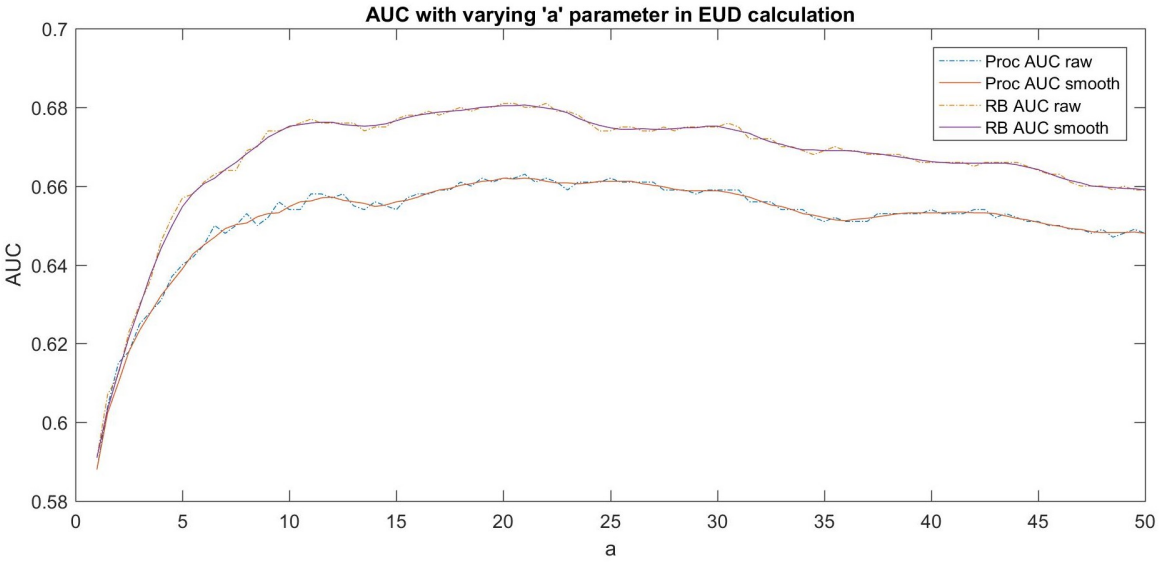


Fig. 3.3 Dose-toxicity sensitivity to EUD parameter a (where $a = 1/n$), shown for CTCAE grade 2 rectal bleeding and CTCAE grade 2 proctitis. Discriminative ability is compared using the area under the receiver operator characteristic curve (AUC).

Table 3.4 Toxicity rates for consolidation (consol.) cohort, training and validation sets, subdivided by prescription level. Faecal incont. = faecal incontinence, Proct. = proctitis, RB = rectal bleeding, GI tox = gastrointestinal toxicity, bowel freq. = bowel frequency, Bother = bowel bother.

Toxicity endpoint	Total consol. n (%)	74 Gy n (%)	60 Gy n (%)	Training n (%)	74 Gy n (%)	60 Gy n (%)	Validation n (%)	74 Gy n (%)	60 Gy n (%)
	186 (100.0)	110 (59.1)	76 (40.9)	139 (75)	82 (59.0)	57 (41.0)	47 (25)	28 (59.6)	19 (40.4)
Diarrhoea	48 (25.8)	30 (16.1)	18 (9.7)	36 (25.9)	22 (15.8)	14 (10.1)	12 (25.5)	8 (17.0)	4 (8.5)
Faecal incont.	32 (17.2)	21 (11.3)	11 (5.9)	24 (17.3)	16 (11.5)	8 (5.8)	8 (17.0)	5 (10.6)	3 (6.4)
Proct. $\geq$ G1	35 (18.8)	18 (9.7)	17 (9.1)	26 (18.7)	13 (9.4)	13 (9.4)	9 (19.1)	5 (10.6)	4 (8.5)
Proct. $\geq$ G2	27 (14.5)	13 (7.0)	14 (7.5)	20 (14.4)	9 (6.5)	11 (7.9)	7 (14.9)	4 (8.5)	3 (6.4)
RB $\geq$ G1	62 (33.3)	32 (17.2)	30 (16.1)	46 (33.1)	24 (17.3)	22 (15.8)	16 (34.0)	8 (17.0)	8 (17.0)
RB $\geq$ G2	21 (11.3)	10 (5.4)	11 (5.9)	16 (11.5)	8 (5.8)	8 (5.8)	5 (10.6)	2 (4.3)	3 (6.4)
GI tox. $\geq$ G1	98 (52.7)	58 (31.2)	41 (22.0)	73 (52.5)	43 (30.9)	30 (21.6)	25 (53.2)	15 (31.9)	10 (21.3)
GI tox. $\geq$ G2	30 (16.1)	15 (8.1)	15 (8.1)	22 (15.8)	11 (7.9)	11 (7.9)	8 (17.0)	4 (8.5)	4 (8.5)
Bowel freq.	45 (24.2)	28 (15.1)	17 (9.1)	34 (24.5)	21 (15.1)	13 (9.4)	11 (23.4)	7 (14.9)	4 (8.5)
Bother $\geq$ G1	73 (39.2)	43 (23.1)	30 (16.1)	55 (39.6)	32 (23.0)	23 (16.5)	18 (38.3)	11 (23.4)	7 (14.9)
Bother $\geq$ G2	43 (23.1)	24 (12.9)	19 (10.2)	32 (23.0)	18 (12.9)	14 (10.1)	11 (23.4)	6 (12.8)	5 (10.6)
Bother $\geq$ G3	17 (9.1)	12 (6.5)	5 (2.7)	13 (9.4)	9 (6.5)	4 (2.9)	4 (8.5)	3 (6.4)	1 (2.1)

3.4 Training and validation data sets

The 186 consolidation cohort patients were split into training (75%, n = 139), and validation (25%, n = 47). The reason for splitting the cohort was to validate final toxicity prediction models using data that were not used in the construction of the model. A model generated using training set data only will have greater variance than models generated using the entire cohort. However, this approach is preferable for strengthening the power of a predictive model according to the Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD) statement [45], and recommendations from experienced lecturers on the ESTRO course on Dose Modelling and Verification for External Beam Radiotherapy held in MAASTRO, 2018. The data were non-randomly split by stratifying patients first by fractionation schedule, then by toxicity incidence. This is considered a Type 2b analysis type in the TRIPOD statement. In order to retain the proportions of toxicity between training and validation sets, patients were randomised for each individual toxicity endpoint. Final toxicity incidence rates for consolidation cohort patients are displayed in Table 3.4 for the entire cohort, as well as training and validation sets, and are further subdivided by prescription level.

3.5 Dose surface maps of the rectal wall

In the early dose-toxicity analysis of the discovery cohort conducted within this thesis, dose to the rectal wall was represented using dose surface maps (DSMs). This was identified as an appropriate analysis technique because the rectum is a hollow tubular-like organ where dose to the rectal wall is the parameter of interest. Furthermore, for the purposes of dose accumulation, DSMs retain spatial information which would be lost if adopting the more conventional dose-volume histogram approach.

The process of constructing the DSMs from the VoxTox dose and structure data was initially developed by Dr Simon Thomas, Head of Medical Physics & Clinical Engineering, Cambridge University Hospitals. Following the methods first described by Buettner et al. [27], a two-dimensional representation of dose to the rectal wall was generated by virtually 'cutting' along the posterior rectal wall and 'unfolding' to produce the dose surface map. Rather than identifying the posterior cutting point as the most posterior point of the rectal contour on a given slice, in VoxTox this point is identified as the point directly posterior to the centroid of the rectal contour. Each slice was normalised to a set number of pixels, defined as the number of slices along the full length of the rectum from the planning CT scan.

3.6 Integrating finite element analysis into VoxTox

For analysis of the consolidation cohort, physical-based biomechanical models were developed for accumulation of dose at the voxel level. Using biomechanical modelling to simulate inter-fraction motion of the rectum provides a more anatomically representative basis for dose accumulation than the 2D DSM approach described previously, which is limited to normalised planar expansion only. Three-dimensional finite element analysis allows high resolution accumulation of dose-histories at the voxel level. Similar to the Morfeus system described previously, the biomechanical model was developed using Abaqus, a Computer Aided Engineering (CAE) Finite Element Analysis (FEA) environment (Dassault Systèmes).

The simulated rectal model was initially developed by Prof Michael Sutcliffe, Cambridge University Engineering Department and was adapted for experimental feasibility studies by BSc and MEng students. Within the scope of this PhD project, the code was modified in order to read and process patient DICOM files, model parameters were optimised for the rectal wall, and the system was integrated into the VoxTox workflow.

Patient data were stored in DICOM proprietary format which requires careful handling and consideration before it can be meaningfully processed. In order to streamline biomechan-

ical modelling of large-scale patient data, an automated workflow was implemented. Total processing time for a patient with 37 fractions was approximately 25 minutes.

An overview of the four data processing stages is shown in Figure 3.4, with detailed explanation below:

3.6.1 Input

The kVCT contours, manually defined for the purposes of treatment planning, and the daily autosegmented MVCT contours were stored as DICOM (RS) structure files which were read into MatLab. The contours were assigned geometric coordinates from the corresponding image set (RI). At the time of treatment, the patient is set up on the couch by aligning small tattoos on the skin with in-room lasers. Image guidance scans (MVCT) were used to confirm that the positioning of the internal anatomy matches that of the treatment planning scan (kVCT). Minor positional corrections can be applied in the right-left, superior-inferior, and anterior-posterior directions via 'couch shifts' to ensure the target is adequately covered (roll can also be corrected for using small gantry rotational offsets). These small daily couch shifts were recorded by the TomoTherapy unit and added to private tags in the DICOM image header during the data extraction process. In order to register all MVCT images, structures, and doses to a common coordinate system, i.e. that of the original kVCT scan, these shifts were appropriately applied.

3.6.2 Pre-processing

Once all patient information was registered, the input file for Abaqus was generated. For each patient, a common starting structure was defined, based upon the centreline of the rectal contour from the kVCT scan. A central mandrel 'innerpart' was generated by expanding from the centreline to a cylindrical tube of fixed radius, with an 'outerpart' consisting of 8-node brick elements. This allows identification and tracking of individual material elements, surrogates for voxels, undergoing simulated deformation throughout the entire course of treatment.

The registration-corrected contours direct the final innerpart positioning using displacement loading, simultaneously expanding the outerpart through surface contact interaction. Deformation of individual elements was dependent upon the biomechanical properties and boundary conditions applied.

3.6.3 Finite element simulation

Biomechanical modelling was run autonomously using input files. Starting from the internal mandrel, the deformation of individual elements depends on the biomechanical properties and boundary conditions applied. Final surface data are extracted from the deformed model using a Python script.

3.6.4 Post-processing

The coordinates of the deformed rectal contour were imported into MatLab. The corresponding DICOM dose (RD) matrix was read and registered to the corrected coordinate system. The delivered dose to each element was then calculated via interpolation. Final delivered dose accounts for the effect of interfraction motion, improving upon previous solutions through the application of biomechanical modelling to predict voxel-level deformations.

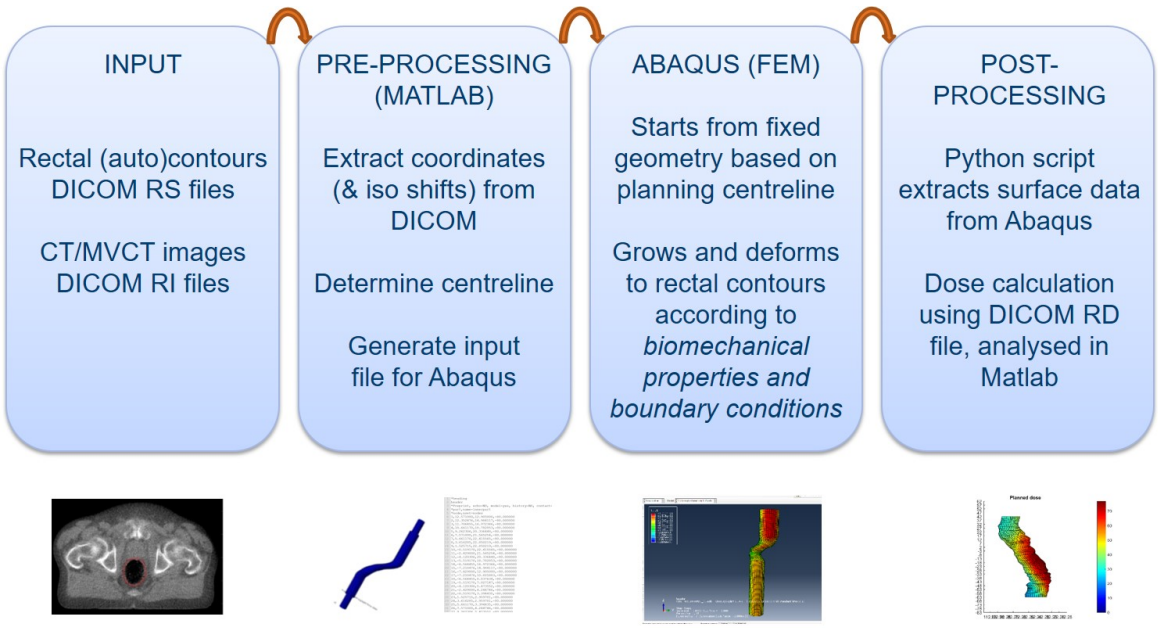


Fig. 3.4 Diagram demonstrating the interaction between patient data (DICOM RS structure, RI image, and RD dose files), finite element modelling (FEM), and final dose output when integrating biomechanical modelling into the VoxTox process

3.7 Summary

The components described in this chapter provide the foundations for the work conducted within this PhD thesis and will be referred to throughout. Automation was a primary focus when developing these systems due to the large scale processing of patient data required. The systems have been collaboratively developed by members of the multidisciplinary VoxTox team with the ultimate aim of addressing the hypothesis that delivered dose to the rectum is a better predictor of toxicity than planned dose in prostate radiotherapy. During this PhD project scientific contributions were made in several areas; identifying where improvements could be made in the autosegmentation software (which led to system updates), comparative measurements were conducted for updates in both the dose calculation and autosegmentation algorithms, dose-toxicity analysis was performed using parameters extracted from DSMs, the biomechanical modelling system was adapted for handling VoxTox patient data, high volumes of toxicity data were processed and analysed, and novel methods were developed for analysing and parameterising accumulated delivered dose to rectum from finite element models, and finally, all data were combined in order to translate and communicate these quantitative measures in terms of risk for developing normal tissue complication during prostate radiotherapy.

Chapter 4

Dose-toxicity analysis using rectal dose surface maps

In this chapter, two-dimensional dose surface maps (DSMs) of the rectal wall are parameterised for both planned and accumulated dose in order to explore dose-toxicity associations. Methods previously developed for quantifying planned DSMs are applied to accumulated DSMs that have been calculated by summation of daily delivered DSMs. Two cohorts from the VoxTox study were investigated: the discovery cohort, collected prior to formalised imaging protocols or collection of baseline data; and the consolidation cohort, with full baseline data and standardised imaging.

4.1 Introduction

In prostate radiotherapy, the correlation between dose to rectum and toxicity has been the focus of many research studies. The rectum is one of the dose-limiting organs when planning intensity-modulated radiotherapy (IMRT) to the prostate due to the risk of radiation-induced adverse effects. Modern systems for inverse IMRT treatment planning iteratively seek to achieve an optimal plan, delivering maximal dose to the tumour volume and minimal dose to healthy organs. Current normal tissue complication probability (NTCP) models and conventional treatment planning constraints are based upon dose–volume histogram (DVH) data to minimise the risk of toxicity. With ever improving disease control [51, 53] and survival rates [32], there is an increasing emphasis on post-treatment quality of life and many research institutes are now focusing on improving the understanding of adverse effects of radiation to healthy organs in order to reduce toxicity incidence.

The DVH-based approach to radiotherapy treatment planning has been criticised for lacking in spatial dose consideration [3]. Consequently, accumulation of DVHs is not dosimetrically representative and results in false overestimations of dose. A review by Landoni et al. [87] emphasises the need to assess associations between spatial dose patterns and late toxicity [47], particularly as results may reveal inhomogeneous intra-organ radiosensitivities.

Several research groups have explored alternative approaches for parameterisation of dose distributions in order to establish links with toxicity. Methods have included dose-surface histograms [26, 27, 74, 121], dose-surface maps [26, 27, 104], dose-line histograms [74], principal component-based pattern analysis [150], and voxel-based approaches for identifying rectal subregions [3, 60]. These studies have been limited in their analysis by the availability of planned dose data only, based on a single anatomical snapshot in time.

A common recommendation in the literature has been the need to establish dose-toxicity models based on delivered dose [82]. However, this has proven technically challenging to date due to hardware and software limitations. These challenges have been addressed within the VoxTox Research Programme [30, 46] where contours generated from on-treatment megavoltage computed tomography (MVCT) image guidance scans are used to calculate daily delivered dose. This approach has made it possible to account for the effect of interfractional anatomical variation. Total delivered dose can be estimated by accumulating daily delivered dose throughout the course of radiotherapy. Studies by the VoxTox group have demonstrated that the rectum moves more than previously predicted based on estimates from prostate motion [136], and that planned dose is not equal to delivered dose [139].

The dose–surface map (DSM) approach has been implemented within this study as a solution enabling meaningful accumulation and conservation of geometric information, an advantage over the DVH methodology. The concept of accumulating DSMs to estimate total delivered dose has been applied previously for the bladder [121]. By extracting spatial parameters from DSMs of delivered dose, and linking with the archive of patient follow-up data available within VoxTox, it was hypothesised that stronger correlations could be established with late toxicity than previously achievable using planned dose alone. Ultimately, improved dose-toxicity modelling based on delivered dose could facilitate real-time in silico prediction of NTCP within the clinical pathway.

4.2 Discovery cohort analysis

4.2.1 Material and methods

4.2.1.1 VoxTox study design & patient information

One hundred and nine prostate cancer patients were selected from the discovery cohort of the VoxTox research programme [30, 46]. This cohort comprised patients treated prior to the formal collection of baseline data, but for whom prospective follow-up data of at least 2 years were available (median 4 years). Early VoxTox patients were selected based on expected benefit from IMRT rather than conventional 3D conformal radiotherapy. Patients in this study were included on the basis of availability of pre-existing toxicity status from clinical notes, or no reported toxicity, and were limited to those prescribed IMRT to a dose of 74 Gy in 37 fractions, the standard of care in the UK at the time [107]. VoxTox patients were treated with TomoTherapy® (Accuray, Sunnyvale, CA). Manual contouring of the anatomy on the kilovoltage computed tomography (kVCT) planning scan was performed according to local procedures [139], adapted from clinical trials. Daily MVCT image guidance scans were acquired immediately prior to treatment delivery for the purposes of online target localisation [29]. Following departmental imaging protocols, scans were inspected for rectal dilation and if deemed excessive, remedial action was taken prior to delivery of radiation therapy [153].

4.2.1.2 Dose–surface map construction & dose accumulation

Within the VoxTox research programme, MVCT scans are multi-functional; primarily for the purpose of routine image guidance, they also provide a platform for calculation of delivered dose. The rectum was identified on each MVCT image series using an in-house autocontouring system based on a customised Chan-Vese segmentation algorithm [143]. Delivered dose was independently calculated using a locally implemented ray-tracing algorithm [152, 154] and the rectal contour-of-the-day, accounting for inter-fraction motion. Automation and integration of dose calculation and contouring systems were essential for large-scale processing of the 4142 scans in this study.

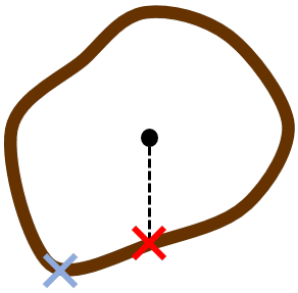


Fig. 4.1 A cutting point directly posterior to the centre of mass of the rectal contour (red cross) at each slice was found to be more stable for DSM dose accumulation than the most posterior point (blue cross).

Planned and daily DSMs were generated based on algorithms described by Buettner et al. [27] and Murray et al. [104]. The rectal wall was considered the structure of interest, and was treated as a

tubular surface rather than a volume. Contours were virtually 'cut' along the superior–inferior axis and 'unfolded' to a two-dimensional plane. The 'cutting point' was identified as the point on the contour surface directly posterior to the centre of mass of the rectal outline, on each CT slice [139], rather than the most posterior point as used by Buettner et al. [27] in their analysis of planned DSMs, as shown in Figure 4.1. For the purposes of accumulating multiple DSMs per patient, the position of the most posterior point was found to be sensitive to daily contour changes and large variations were observed. The cutting point selected here, posterior to the rectal contour centre of mass, was found to be a more stable landmark.

The length of the planned-DSM was defined by the number of slices of the manually contoured rectum on the kV planning scan (slice thickness 3 mm). The circumference of the rectal contour on each slice was normalised such that the unfolded width of the planned-DSM was equal to the length. Daily delivered DSMs calculated from the image-guidance MVCT scans (slice thickness 6 mm) were normalised to the same width as the planned-DSM but were restricted in length by the field of view (FOV), resulting in a shorter DSM, as shown in Figure 4.2.

Rectal DSMs were calculated for each treatment fraction, and corrected for daily couch shifts. For the purposes of dose accumulation, any 'missing' dose data cropped by the restricted FOV superiorly or inferiorly (as indicated in grey in Figure 4.2) were substituted from the planned-DSM [139] in order to maintain common dimensions between final accumulated-DSM and planned-DSM . The final accumulated-DSM was resampled to match the 3mm resolution of the planned-DSM, producing an easily comparable and interpretable spatial

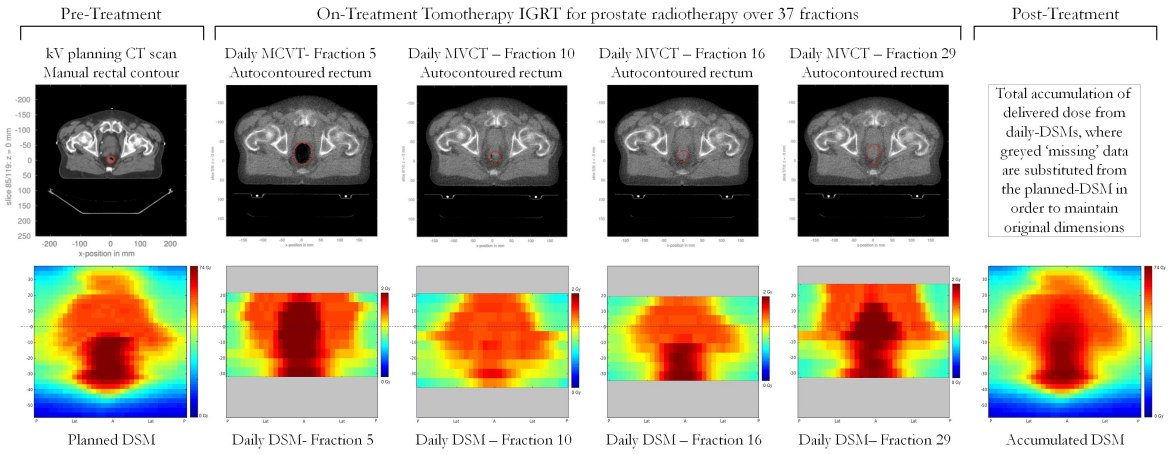


Fig. 4.2 Generation of planned, daily delivered, and total accumulated dose surface maps (DSMs). Four of 37 daily DSMs are displayed. Daily DSMs are limited in length by the imaging field-of-view, and 'missing' (greyed) data were substituted using planned dose for the purposes of total dose accumulation.

representation of total delivered dose to the rectal wall throughout the course of prostate radiotherapy [154].

The use of planned-DSM data as a surrogate beyond the boundaries of the MVCT FOV is assumed to be an acceptable estimate because the relative anatomical motion of the rectum becomes more confined by surrounding musculature as the distance from the prostate increases in the inferior direction [139]. However, this could reduce detection sensitivity of potential differences and is a limitation when using DSMs to compare planned and accumulated dose.

4.2.1.3 Dose parameters & clinical endpoints

Dose was parameterised from DSMs using two methods implemented in MATLAB ®(MathWorks, Natick, MA):

1. Calculation of Equivalent Uniform Dose (EUD)

2. Fitting of DSM 'dose-widths' to discrete isodose clusters

Equivalent Uniform Dose (EUD) reduces the dose information extracted from the DSMs to a single generalised value which allows comparison between inhomogeneous dose distributions [113]. An 'a' value of 11.11 was used to calculate EUD [94], as discussed in Section 3.3. Spatial dose information was generated by reproducing Buettner's ellipse-fitting method [27], reporting the most significant dose quantifier, the lateral extent, termed here the 'DSM dose-width'. An example case demonstrating how DSM dose-widths were extracted is shown in Figure 4.3.

For a given isodose level, a binary image was created from the DSM by assigning a pixel value of 1 to doses greater than or equal to the nominated isodose, with lower doses assigned a value of 0. An ellipse was then fitted to the largest central cluster. The maximum lateral extent of the ellipse was projected onto the DSM axis, accounting for any rotation with respect to the DSM coordinate system. The resulting DSM dose-width, expressed as a the percentage of total normalised DSM width, allowed parameterisation of the geometrical dose distribution which would have been masked using a DVH approach.

For each patient, EUD and DSM dose-widths for isodose levels of 30, 40, 50, 60, 65, and 70 Gy were calculated from planned-DSM and accumulated-DSM. Doses less than 30 Gy were not included as DSM dose-width measurements became dominated by values greater than 100 %, indicating that the entire rectal circumference was receiving more than or equal to the selected isodose level. By comparison to the binary image shown in Figure 4.3, this would appear as a white band spanning the width of the DSM, at the plane where the

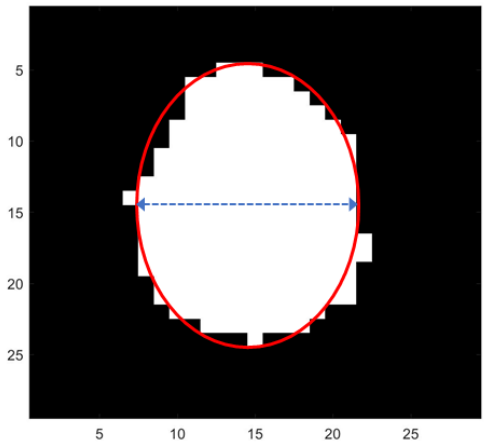


Fig. 4.3 Calculation of DSM dose-widths. Binary images were generated in MATLAB at each discrete dose level. The area in white indicates the largest cluster of pixels receiving a dose greater than or equal to the specified dose level. MATLAB's *regionprops* tool was used to fit an ellipse to the isodose cluster (red) and extract the width of that ellipse (blue dashed arrow) as the DSM dose-width. This example shows the isodose cluster and ellipse fitted to a dose level of 30 Gy for a patient prescribed 60 Gy in 20 fractions. The measured DSM-dose width was 53 %.

DSM dose-width was being measured. Hence, DSM dose-widths at low dose levels were being extrapolated beyond the dimensions of the DSM, and were therefore excluded from the analysis. This was identified as a limitation of the ellipse fitting method when seeking to analyse low dose toxicity correlations. At the other end of the dose scale, doses greater than 70 Gy were also excluded from toxicity analyses due to the increasing frequency of 0 % DSM dose-widths, indicating that doses greater than or equal to the selected isodose level were not received by the rectal wall. By comparison to Figure 4.3, the binary image would be entirely black at the highest dose levels. From accumulated dose at 74 Gy, 55 % of patients recorded a 0 % dose-width, increasing to 91 % at 75 Gy (compared with 3 % and 41 % for planned dose, respectively). It was identified that a 0 % DSM dose-width result could conceal information leading to misinterpretation of data when performing AUC calculations so results at these dose levels were not reported. Despite these restriction, the dose levels included within this study incorporate the 39–61 Gy range at which Buettner [27] determined significant correlations between lateral extent and toxicity.

Study specific clinical reporting forms were developed by clinical colleagues when setting up VoxTox to ensure robust collection of toxicity data, and raw data were used to populate recognised systems, including: Common Terminology Criteria for Adverse Events (CTCAE) v4.03 [158], Late Effects of Normal Tissues/Subjective, Objective, Management, Analytic (LENT SOMA) scores [1]; Radiation Therapy Oncology Group (RTOG) grading

Table 4.1 Clinical endpoints, scoring systems and incidence rates of the 6 most frequently reported toxicities within the patient sample. † LENT SOMA [1] rectal bleeding ≥ grade 2 was equivalent to CTCAE [158] rectal bleeding ≥ grade 1. * Data were missing for 4 patients so sample size was reduced to 105.

Clinical Endpoint	Scoring System	Incidence% (n)
Objective rectal bleeding ≥ grade 2†	LENT SOMA [1]	25.7 (28/109)
Proctitis ≥ Grade 2	Gulliford [70]/ RTOG [127]	16.5 (18/109)
Sphincter control ≥ grade 1	LENT SOMA [1]	10.1 (11/109)
Subjective rectal pain ≥ grade 1	LENT SOMA [1]/ CTCAE [158]	15.6 (17/109)
Bowel problems ≥ grade 1	Gulliford [70]	30.7 (32/105*)
Bowel problems ≥ grade 2	Gulliford [70]	11.5 (12/105*)

system [127]; University of California, Los Angeles, Prostate Cancer Index (UCLA-PCI) questionnaire [89]. Receiver Operator Characteristic (ROC) curves were generated using SPSS®(IBM®23.0.0.2) to evaluate the link between dosimetric parameters extracted from planned and accumulated DSMs, and the six most prevalent clinical endpoints, listed in Table 4.1.

The mean area under the curve (AUC), with associated upper and lower 95% confidence intervals (CIs), was calculated for each ROC curve as a measure of the level of association between dosimetric parameter and toxicity. An ideal correlation would have an AUC of 1. Results were reported for dosimetric parameters with AUC≥0.6 and lower 95% CI≥0.5, considered statistically significant by Gulliford et al. [71]. Generally, an AUC greater than 0.6 indicates acceptable discrimination, with AUCs below this level considered poor or non-discriminative.

4.2.2 Results

4.2.2.1 Difference between planned and accumulated dose

Accumulated delivered dose was found to be systematically lower than planned dose. For all patients, EUD of accumulated-DSM was lower than that of planned-DSM (mean difference -2.2 Gy, standard error 0.3 Gy, range [-7.1, -0.3] Gy). Table 4.2 shows EUD breakdown for all patients, as well as comparing patients with and without toxicity. DSMs provide insight into the spatial distributions of regions of high dose difference. Difference maps between planned and accumulated DSMs (Figure 4.4a) indicate concentrated regions of lower accumulated dose up to -4 Gy corresponding to right and left lateral sides of the middle-third of the rectal wall. Lower magnitudes of dose difference were observed in the mid-anterior rectal wall

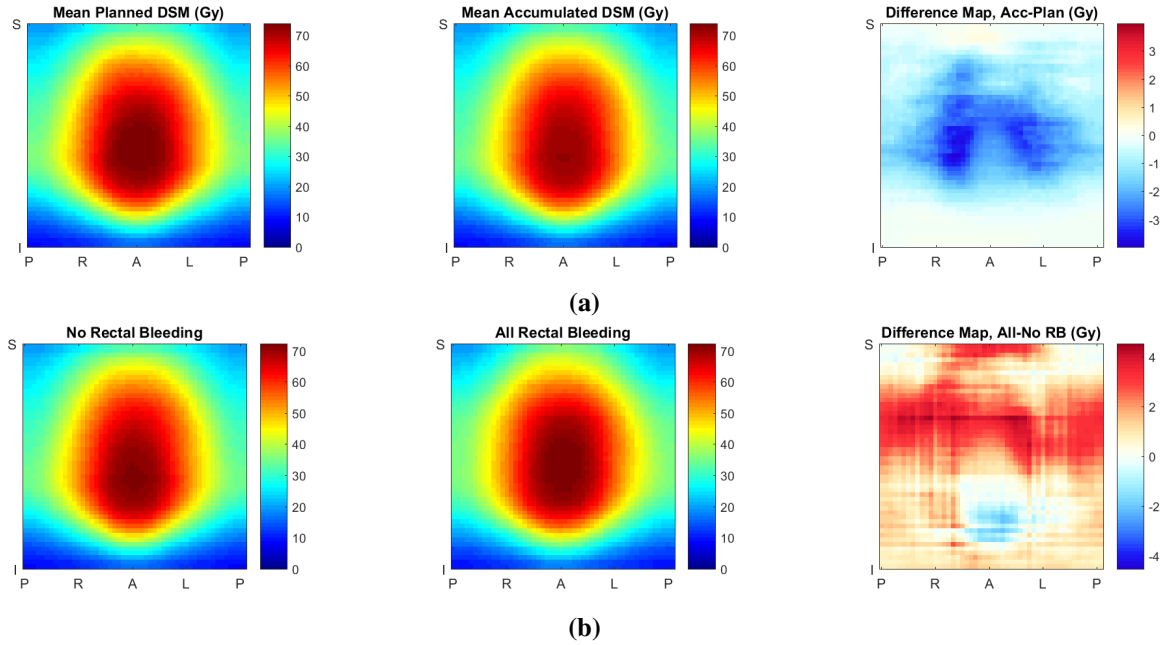


Fig. 4.4 (a) Difference between mean planned dose surface maps (DSM) and mean accumulated DSM for 109 patients from the discovery cohort, (b) Mean accumulated dose surface maps for patients with (n = 28) and without (n = 81) rectal bleeding (RB). S = superior, I = inferior, P = posterior, R = right, L = left, A = anterior.

where the highest dose levels are recorded, closest to the prostate. This suggests that dose to the right and left lateral sides are most affected by interfraction motion of the rectum. This information is lost if considering EUD as a single spatial metric alone.

Splitting 109 patients from the discovery cohort into groups with and without rectal bleeding, and calculating the mean DSM per group reveals further information (Figure 4.4b). Results show that those experiencing RB received a greater dose than those who did not, and the spatial DSMs in Figure 4.4b indicate that on average, these patients received a higher dose of over 4 Gy across a lateral band spanning the circumference of the mid to upper rectal wall.

4.2.2.2 Linking dosimetric parameters with rectal toxicity

Rectal bleeding

Twenty-eight patients reported rectal bleeding CTCAE $\geq$ Grade 1, which was equivalent to LENT SOMA $\geq$ Grade 2. The AUC was greater for all accumulated DSM dose-widths than planned DSM dose-widths up to 70 Gy (Table 4.3). Accumulated dose was found to be discriminative of rectal bleeding at three DSM dose-widths where planned dose was not; 60 Gy, 30 Gy, and 40 Gy, ordered greatest to least significance (see Table 4.3). As shown in Figure 4.5 (a), the lower 95% CI for the corresponding planned DSM dose widths extended

Table 4.2 Breakdown of equivalent uniform dose (EUD) calculation for mean planned and accumulated dose surface maps (DSMs), split by patients experiencing CTCAE rectal bleeding (RB) for 109 patients from the discovery cohort. For toxicity comparisons, values in brackets indicate difference with respect to No RB group.

	Mean EUD (Gy) Planned DSM	Mean EUD (Gy) Accumulated DSM	Difference (Gy)
All patients	60.06	57.99	-2.07
No RB (n=81)	59.82	57.70	-2.12
RB $\geq$ G1 (n=28)	60.84 (+1.02)	58.92 (+1.22)	-1.92
G1 RB (n=18)	60.34 (+0.52)	58.09 (+0.39)	-2.24
G2 RB (n=9)	62.00 (+2.18)	60.77 (+3.07)	-1.23

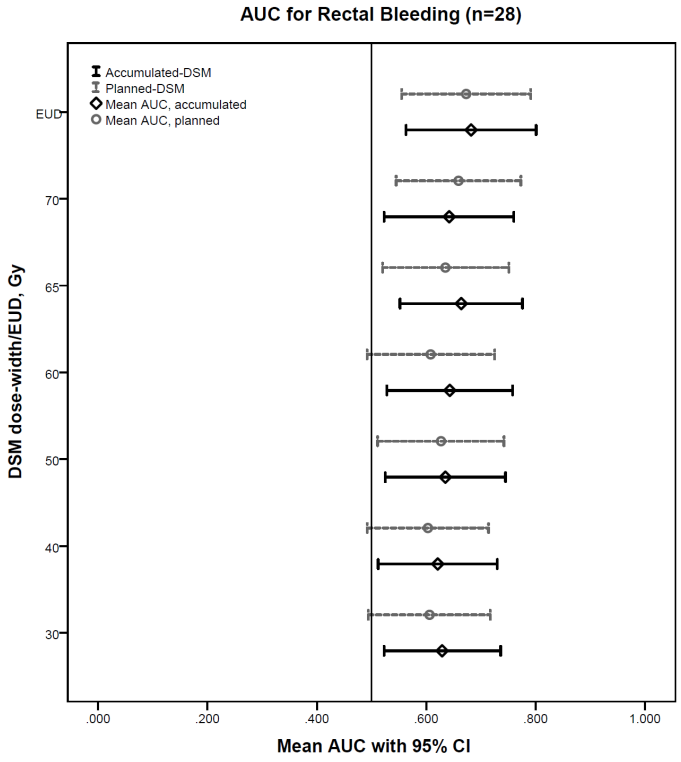
below 0.5 indicating that planned dose can not be considered a significant predictor of rectal bleeding at these DSM dose-width levels.

A similar level of statistical correlation was observed for both accumulated and planned dose with rectal bleeding at three DSM dose-widths; 50, 65, and 70 Gy. The strongest spatial predictor of rectal bleeding was the accumulated 65 Gy DSM dose-width (AUC 0.664), and the largest difference between planned and accumulated DSM dose widths was at 60 Gy (AUC difference 0.035). The relatively low variability between different AUCs at different dose levels may be due to the collinear nature of radiotherapy dose delivery, where a greater high-dose proportion subsequently leads to a greater low-dose contribution depending on the gradient of dose fall-off.

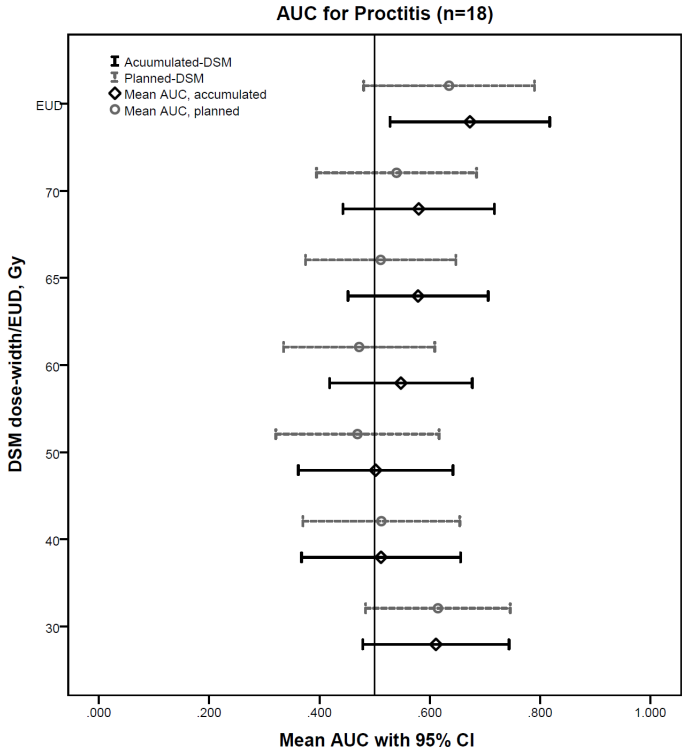
Overall, EUD exhibited the strongest correlation with rectal bleeding for both accumulated and planned dose, AUC 0.682 and 0.673 respectively.

Table 4.3 Mean Area Under the Curve (AUC) for planned and accumulated DSM dose-widths and EUD corresponding to rectal bleeding $\geq$ Grade 2 (LENT SOMA) and $\geq$ Grade 1 (CTCAE), n = 28/109. The greater AUC of each parameter has been presented in bold.

Dose Level (Gy)	Mean AUC (Planned)	Mean AUC (Accumulated)
30	0.606	**0.629**
40	0.603	**0.621**
50	0.627	**0.635**
60	0.608	**0.643**
65	0.635	**0.664**
70	**0.659**	0.642
EUD	0.673	**0.682**



(a)



(b)

Fig. 4.5 Mean area under the receiver operator characteristic curve (AUC) for dosimetric predictors of (a) rectal bleeding and (b) proctitis. Error bars indicate 95 % confidence intervals (CI).

Proctitis

Eighteen patients reported RTOG proctitis $\geq$ Grade 2. Accumulated-EUD was found to be the sole significant result correlating with RTOG/Gulliford proctitis $\geq$ Grade 1 (AUC 0.673) (Figure 4.5(b)). Planned EUD was not considered to be a significant indicator as the lower 95% CI $\leq$ 0.5. Accumulated DSM dose-widths had greater AUC than planned DSM dose-widths at 50, 60, 65 and 70 Gy, and were equivalent at 40 Gy. At 30 Gy, the AUC of planned DSM dose-width was slightly higher than the accumulated AUC (0.004 difference). Differences between planned and accumulated dose were greater than observed for rectal bleeding, and the variability in mean AUC was greater between dose levels, perhaps supporting the suggestion that different endpoints have different pathophysiologies, which in turn may have different radiosensitivities.

4.2.3 Discussion

Radiation dose received by the rectal wall during prostate radiotherapy was calculated and accumulated using DSMs. Geometric aspects of dose distribution - information not distinguishable from DVHs - were parameterised using DSM dose-widths. EUD was calculated to compare planned and accumulated DSMs using a single metric. Extracted dosimetric parameters were evaluated against six clinical endpoints reported by patients within the VoxTox research programme. Previous dose-toxicity investigations in the literature have been limited to planned dose only. This study has demonstrated, for the first time, that delivered dose can be a stronger predictor of toxicity in the case of rectal bleeding and proctitis in prostate radiotherapy.

Toxicity rates reported in the literature have been variable. The rate of bowel toxicity $\geq$ Grade 2 (CTCAE), 5 year cumulative incidence, amongst VoxTox prostate patients was 17%. This falls within the bowel toxicity $\geq$ Grade 2 (RTOG) range of 13.7% - 24.9% for IMRT over the same timeframe, reported by Dearnaley et. al [52] and Wortel et. al [170], respectively. The rates of incidence indicate that toxicity remains an important clinical issue.

Many associations were found between DSM dose-widths with rectal bleeding. Accumulated DSMs generated greater AUCs than planned DSMs for 5 DSM dose-width levels up to 70 Gy. The strongest correlation between rectal bleeding and any spatial parameter was the 65 Gy DSM dose-width from accumulated dose (AUC 0.664). At 30, 40 and 60 Gy, accumulated DSM dose-widths produced AUC$\geq$ 0.6 and lower 95% CI$\geq$ 0.5, where corresponding planned DSM dose-widths did not. These thresholds were considered indicative of significance following the methods of Gulliford et al [71]. The greatest difference between planned and accumulated AUCs was observed at the 60 Gy DSM dose-width. Overall,

the results compared well with the findings of Buettner et al [27] who reported the most significant correlation with rectal bleeding to be the 61 Gy lateral extent (AUC 0.66), derived from planned dose data.

Accumulated EUD was found to have the strongest correlation overall with rectal bleeding (AUC 0.682), and was the only predictor of proctitis (AUC 0.673).

For all patients, accumulated-EUD was systematically lower than planned-EUD. A contributory factor was possibly the inherent blurring of high dose regions during accumulation. Upon visual inspection of daily DSMs, the differences in size, shape and position of the high dose region due to anatomical variation was clearly visible (for example, shown in deep red in Figure 4.2). During accumulation, high doses were superimposed in overlap regions, but reduced where isodose edges differed, due to averaging over the full course of radiotherapy. This affected the maximum dose of the final accumulated-DSM, on which EUD calculation was heavily weighted.

The dose-blurring effect could also have been responsible for the increased frequency of 0% DSM dose-width results at high dose levels from accumulated-DSMs with respect to planned-DSMs. At 70 Gy, 4/109 patients recorded a 0% accumulated DSM dose-width (including 1 patient experiencing toxicity), whereas all corresponding planned DSM dose-widths had non-zero results. Furthermore, dose levels could not be considered independent variables, as a low 70 Gy DSM dose-width was likely to be associated with a low 65 Gy DSM dose-width, and a cooler plan overall. These issues were not accounted for within the scope of this study.

The generally lower reported values for EUD and DSM dose-widths from accumulated dose compared with planned dose should not be interpreted as delivered treatment erring on the ‘safe side’ in terms of dose to rectum. Where current NTCP models are based on planned dose, the presented results suggest that the same magnitude of risk would be associated with a systematically lower delivered dose.

The findings for 109 patients from the discovery cohort show that the difference in dose between patients with and without rectal toxicity is greater from delivered dose than planned dose. This indicates that dosimetric parameters from accumulated-DSMs could provide new information to improve understanding of the relationship between dose and toxicity. The single parameter EUD was a superior predictor of rectal bleeding and proctitis than spatial dose quantifiers. However, DSM-dose widths produced several strong correlations with rectal bleeding, and for 5/6 dose levels, accumulated dose generated AUC values greater than planned dose.

Parameterisation of delivered dose to the rectal wall during prostate radiotherapy has revealed stronger correlations with rectal bleeding and proctitis than achievable from planned

dose. New information from accumulated delivered dose could lead to improved dose-toxicity modelling in the future, with the aim of reducing post-treatment toxicity.

4.3 Consolidation cohort analysis

4.3.1 Methods and materials

The consolidation cohort includes patients who were prospectively identified to be enrolled into the VoxTox study once the programme had been granted ethics approval in 2013. Formalised contouring guidelines, imaging protocols, baseline data collection, and regular follow-up questionnaires were formally implemented for this cohort [30, 137].

Recruitment to the VoxTox study closed in March 2018, so follow-up data continue to be collected and mature for a large number of patients. The patient database was frozen in January 2018, allowing interim analysis of the consolidation cohort for the purposes of this research project. A summary of the frozen cohort is presented in Table 4.4.

Table 4.4 VoxTox prostate patient figures, January 2018. * One patient was removed from the analysis due to incomplete data, leaving n = 186.

	60 Gy	74 Gy	Total
Total number of patients	161	326	508
Number of consolidation patients	111	134	245
Consol. patients with 2 year follow-up	77	110	187*

The consolidation cohort (n = 186) were split into training (75 %, n = 139) and validation (25 %, n = 47) sets. For each endpoint (Table 4.5), the incidence rates of patient toxicity were conserved, as well as the relative proportions of patients prescribed 60 Gy and 74 Gy. After dividing patients into groups to ensure toxicity and prescription parity was maintained, patients were randomised to either the training or validation set. This resulted in a different training set composition for each of the 12 endpoints, as shown in Table 4.5.

This investigation was limited to the training set only. During subsequent analysis, one patient was removed from the training set due to a failing finite element model, and was retrospectively removed from the analysis conducted in this chapter. This has been reflected in the patient numbers and analysis presented here.

The toxicity endpoints investigated were: CTCAE diarrhoea $\geq$ G1 [158], CTCAE faecal incontinence $\geq$ G1 [158], CTCAE proctitis $\geq$ G1 and $\geq$ G2 [158], CTCAE rectal bleeding $\geq$ G1 and $\geq$ G2 [158], RTOG gastrointestinal (GI) toxicity $\geq$ G1 and $\geq$ G2 [127], stool frequency $\geq$ G1 [1], and UCLA PCI bowel bother $\geq$ G1, $\geq$ G2, and $\geq$ G3 [89]. Incidence

Table 4.5 Consolidation cohort training set split by toxicity endpoint and prescription group

	Toxicity rates in training set (n = 139)		Toxicity rates in 74 Gy group (n = 82)		Toxicity rates in 60 Gy group (n = 57)	
Toxicity endpoint	n (tox)	(%)	n (tox)	(%)	n (tox)	(%)
Diarrhoea	36	25.9	22	26.8	14	24.6
Faecal incontinence	24	17.3	16	19.5	8	14.0
Proctitis $\geq$ G1	26	18.7	13	15.9	13	22.8
Proctitis $\geq$ G2	20	14.4	9	11.0	11	19.3
Rectal bleeding $\geq$ G1	46	3.1	24	29.3	22	38.6
Rectal bleeding $\geq$ G2	16	11.5	8	9.8	8	14.0
RTOG GI toxicity $\geq$ G1	73	52.5	43	52.4	31	54.4
RTOG GI toxicity $\geq$ G2	22	15.8	11	13.4	11	19.3
Bowel frequency	34	24.5	21	25.6	13	22.8
Bowel bother $\geq$ G1	54	38.8	32	39.0	22	38.6
Bowel bother $\geq$ G2	31	22.3	18	22.0	13	22.8
Bowel bother $\geq$ G3	13	9.4	9	11.0	4	7.0

rates are shown in Table 4.5.Endpoints were included on the basis of previous endpoints reported in the literature [70, 77]. When splitting into training and validation cohorts, the rates of toxicity incidence were maintained as closely as possible for each prescription group within each cohort. As a result, discrete training sets were randomised per endpoint. Further details have been discussed in Chapter 3, Section 3.4. To combine patient prescription groups, the DSMs of the 60 Gy group were converted to the equivalent dose in 37 fractions (EQD37) using an α/β ratio of 2.149 (as derived in Chapter 3, Section 3.3). The EQD37 correction was applied to each pixel of the planned and final accumulated delivered dose surface maps (DSMs) per applicable patient. A breakdown of the training cohort by toxicity endpoint and prescription group is included in Table 4.5.

As in Section 4.2, dose surface maps (DSMs) of the rectal wall were generated by virtually cutting the rectum at the point directly posterior to the centroid of the rectal contour, at each slice. The width of the rectal DSM was set to be equal to the length of the rectum as defined on the manually delineated kVCT planning scan. For each slice, the rectal circumference was isotropically normalised to the DSM width, where the width remains constant per patient. For the daily MVCT IGRT images where the field of view is restricted, the width remains fixed, but where the field of view is shorter than the full length of the rectum in the superior and/or inferior regions, the 'missing' data are substituted with kVCT planned data as a best estimate (as shown in Figure 4.2).

For all patients, the dosimetric parameters extracted from the planned and accumulated DSMs were equivalent uniform dose (EUD) and dose-widths from 30 - 70 Gy in 5 Gy

increments. Where a 0 % dose-width was recorded, the measurement was removed from the analysis for the dose-width under investigation, alongside the corresponding planned or accumulated dose measurement for the particular patient. This was because a 0 % dose-width could conceal information leading to a misrepresentation of the data [142]. As a consequence, across all endpoints, 1 - 2 measurements were removed from the 65 Gy dose-width, and 6 - 10 measurements were removed from the 70 Gy dose-width. This was not expected to significantly affect discrimination AUCs with respect to lower dose levels, but may have led to wider confidence intervals.

Dosimetric parameters were evaluated for their discriminative ability of predicting each endpoint using the area under the curve (AUC) calculated from the receiver operator characteristic (ROC) curve. Differences between planned and delivered DSMs were analysed using the rectal bleeding validation cohort as a representative sample.

4.3.2 Results

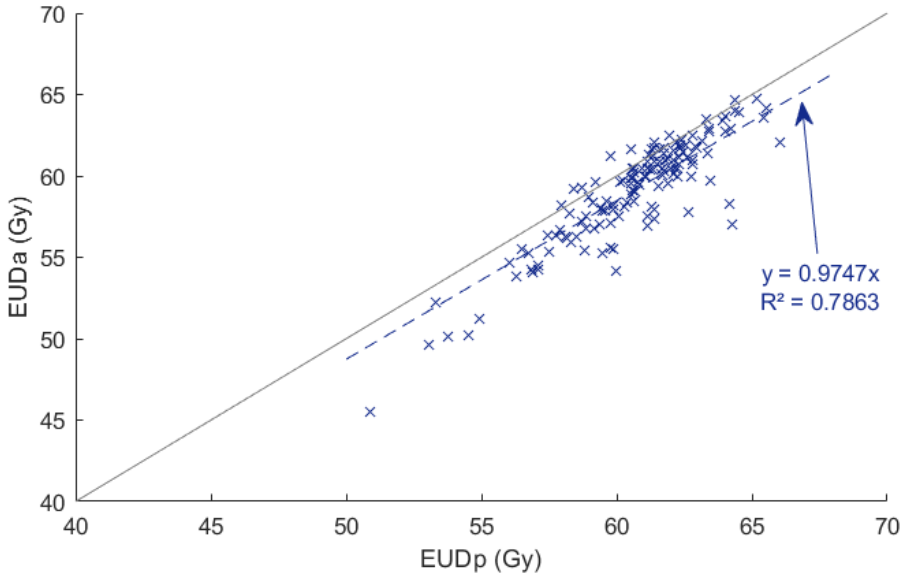
4.3.2.1 Planned versus accumulated dose

Planned EUD was greater than accumulated EUD for 125 of 139 patients, as shown in Figure 4.6. The mean difference in EUD was -1.5 Gy, SE 0.1 Gy, range [-7.2, 1.4] Gy. For 89 patients (63 %), the difference was greater than ± 1 Gy. The mean absolute difference between planned and accumulated dose-widths increased with increasing dose level (Figure 4.7). Mean planned dose-widths were greater than mean accumulated dose-widths at dose levels above 30 Gy, with a maximum difference of -5.2 % at the 70 Gy dose-width.

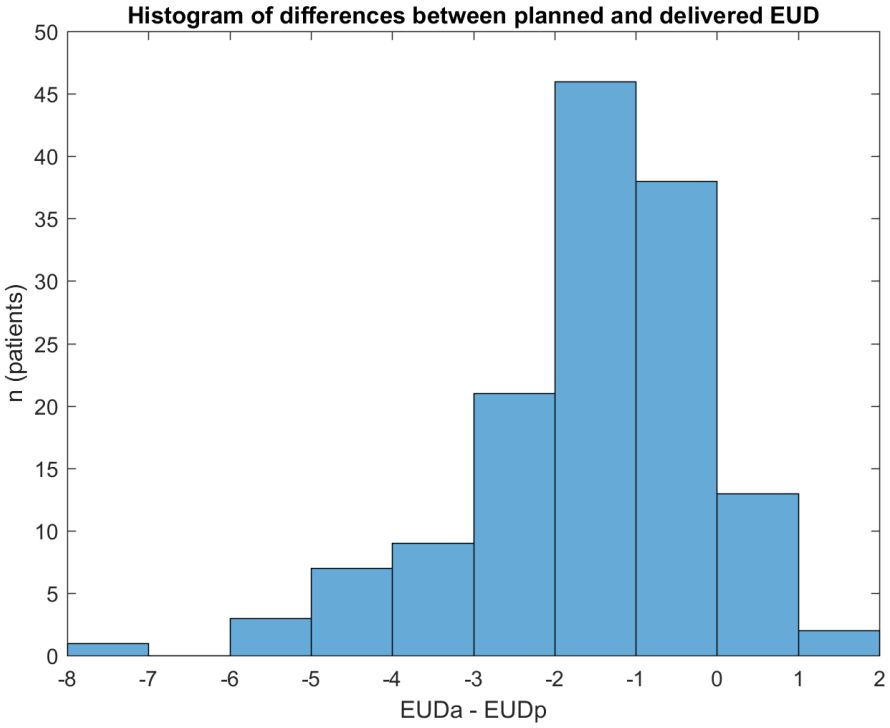
Each planned dose parameter was very strongly positively correlated with the corresponding accumulated dose parameter ($r > 0.8$, $p < 0.001$), indicating that an increase in planned dose results in an increase in delivered dose. The results of a paired samples t-test showed that average differences between planned and accumulated dose were significant for EUD ($p < 0.001$), and dose-widths 40 Gy ($p = 0.002$), and 45 - 70 Gy ($p < 0.001$).

4.3.2.2 Dose-toxicity associations

Results of dose-toxicity analysis for all endpoints are discussed below. The highest performing dose metrics are listed in Table 4.6. The area under the receiver operator characteristic curve (AUC) and 95 % confidence intervals (95%CI) for rectal bleeding $\geq$ G1 and proctitis $\geq$ G2 are shown in Figure 4.8 for comparison with the previous discovery cohort analysis shown in Figure 4.5. AUC indicates the ability of the dosimetric parameter to correctly discriminate between patients with and without toxicity. An ideal correlation would have an AUC of 1. Results were considered significant if AUC ≥ 0.6 and lower 95%CI ≥ 0.5 [71].



(a)



(b)

Fig. 4.6 (a) Comparison between EUD calculated from planned and accumulated dose surface maps (EUDp and EUDa, respectivly). Grey line indicates y = x reference. Dashed line indicates line of best fit, where equation assumes intercept at the origin, and R^2 is an indicator of the goodness-of-fit. (b) Histogram of differences between planned and accumulated EUD.

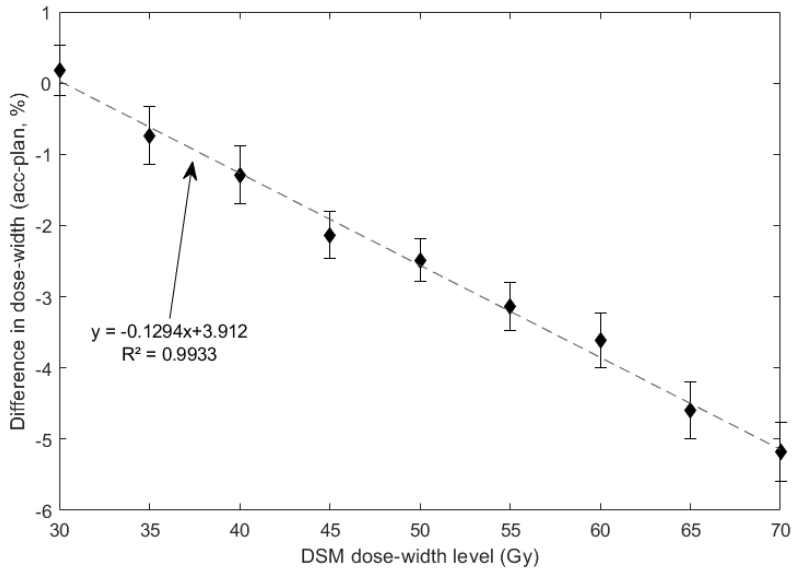


Fig. 4.7 Difference between planned (plan) and accumulated (acc) DSM dose-widths. Dose difference was found to increase with dose level. Dotted line shows line of best fit. Error bars indicate standard error of the difference.

Diarrhoea (n = 36, 25.9%)

The strongest predictor of diarrhoea was the planned 35 Gy dose-width, with AUC = 0.689 [95%CI: 0.587, 0.792]. Planned dose-widths of 30-50 Gy were also predictive. At these dose-widths, planned AUCs were greater than corresponding accumulated AUCs, which were also significant. Additionally, the accumulated 55 and 60 Gy dose-widths and EUD were predictive of diarrhoea, where corresponding planned dose parameters were not.

Faecal incontinence (n = 24, 17.3%)

Planned and accumulated dose-widths at 40 - 60 Gy achieved AUCs greater than 0.6. However, all lower bounds of the 95%CI were less than 0.5, so no single dosimetric parameter could be considered significantly predictive of faecal incontinence.

Proctitis $\geq$ G1 (n = 26, 18.7%)

Planned dose-widths at 45 - 65 Gy and EUD were predictive of G1 proctitis, with the 65 Gy dose-width producing the greatest AUC (0.656 [0.554, 0.757]). The accumulated 60 Gy dose-width and EUD were also significant predictors, but less so than the corresponding planned dose metrics.

Table 4.6 Dosimetric parameters most discriminative of toxicity from planned and accumulated dose surface maps. Dose-width levels or equivalent uniform dose (EUD) resulting in the greatest area under the receiver operator characteristic curve (AUC) are listed. Results for faecal incontinence have been omitted as no parameter produced an AUC with minimum 95% confidence interval greater than 0.5, so was not considered significant. Dashed entry indicates no significant parameter was determined.

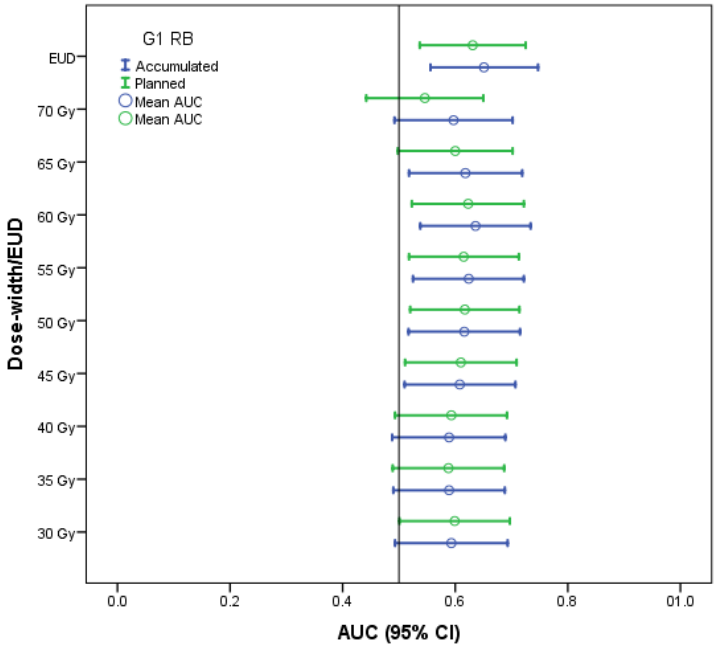
	Planned dose		Accumulated dose	
Toxicity endpoint	Dose-width (Gy)/EUD	AUC	Dose-width (Gy)/EUD	AUC
Diarrhoea	35	**0.689**	35	0.642
Proctitis $\geq$ G1	65	**0.656**	EUD	0.626
Proctitis $\geq$ G2	60	0.635	60	**0.648**
Rectal bleeding $\geq$ G1	EUD	0.631	EUD	**0.651**
Rectal bleeding $\geq$ G2	65	**0.796**	60	0.788
RTOG GI toxicity $\geq$ G1	-	-	EUD	**0.604**
RTOG GI toxicity $\geq$ G2	45	0.722	50	**0.743**
Stool frequency	35	0.646	45	**0.648**
Bowel bother $\geq$ G1	40	0.607	60	**0.638**
Bowel bother $\geq$ G2	-	-	EUD	**0.648**
Bowel bother $\geq$ G3	-	-	EUD	**0.686**

Proctitis $\geq$ G2 (n = 20, 14.4%)

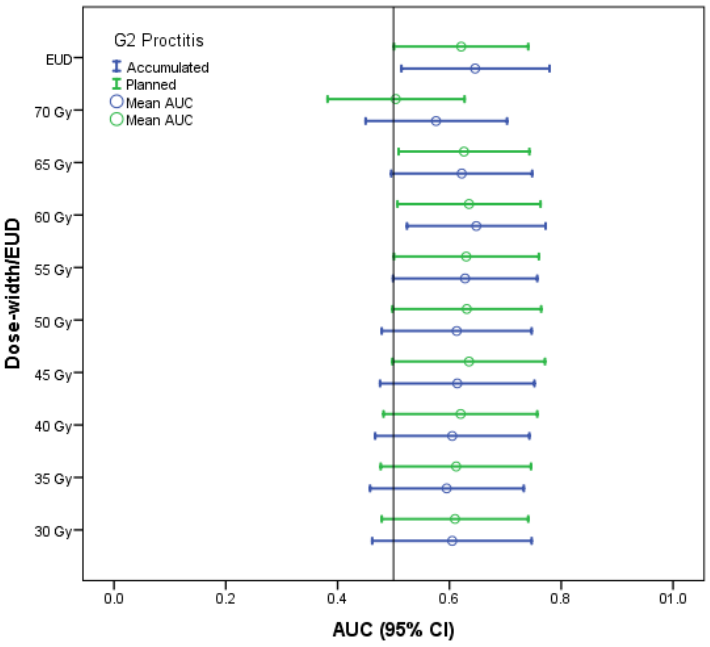
The strongest predictor of proctitis G2 was the accumulated 60 Gy dose-width (AUC 0.648 [0.524, 0.772]), as shown in Figure 4.8b. This differs from the discovery cohort result, which found the only association with G2 proctitis to be accumulated EUD. However, accumulated EUD was also predictive of G2 proctitis for the consolidation cohort analysed here, along with planned dose-widths at 55-65 Gy and EUD. The AUCs from all accumulated dose parameters increased with respect to those reported for G1 proctitis. Of the significant planned dose parameters, the AUCs reduced with respect to those reported for G1 proctitis, with the exception of the 55 Gy dose-width (increase in AUC by 0.001, but wider 95%CI).

Rectal bleeding $\geq$ G1 (n = 46, 33.1%)

Accumulated EUD was the strongest predictor of G1 rectal bleeding (AUC 0.651 [0.556, 0.747]), as shown in Figure 4.8a, supporting the findings of the discovery cohort analysis in Section 4.2. Accumulated 45-65 Gy dose-widths were also associated with G1 rectal bleeding, and 60 and 65 Gy dose-widths were greater than corresponding planned metrics. Dosimetric parameters from planned dose predictive of rectal bleeding G1 were EUD and 45 - 60 Gy dose-widths.



(a)



(b)

Fig. 4.8 Consolidation cohort dose-toxicity analysis. Area under the receiver operator characteristic curve (AUC) is shown for dose-widths and equivalent uniform dose (EUD), calculated from planned and accumulated dose surface maps (DSMs) of the rectal wall for (a) Rectal Bleeding $\geq$ grade 1, and (b) Proctitis $\geq$ grade 2. Results were considered significant where mean AUC > 0.6 and lower 95% CI > 0.5.

Rectal bleeding $\geq$ G2 (n = 16, 11.5%)

The strongest predictor of G2 rectal bleeding was the planned 65 Gy dose-width, with AUC 0.796 [0.705, 0.887]. This was the highest AUC recorded across all toxicity endpoints. All dosimetric parameters were associated with G2 rectal bleeding, with the exception of the planned 70 Gy dose-width. Accumulated AUCs were greater than corresponding planned AUCs for EUD, 30, 35, 55, and 70 Gy dose-widths. Planned AUCs were greater than accumulated for 40, 45, 60, and 65 Gy dose-widths.

RTOG GI toxicity $\geq$ G1 (n = 73, 52.5%)

The only dosimetric parameter associated with G1 GI toxicity was accumulated EUD (AUC 0.604 [0.510, 0.698]). No other parameter produced an AUC $\geq$ 0.6 or lower 95%CI $\geq$ 0.5.

RTOG GI toxicity $\geq$ G2 (n = 22, 15.8%)

The accumulated 50 Gy dose-width was the strongest predictor of G2 GI toxicity (AUC 0.743 [0.621, 0.865]). All AUCs were greater for G2 GI toxicity than G1. Dose-widths of 30-65 Gy were associated with G2 GI toxicity for both planned and accumulated DSMs. Accumulated dose-widths had greater AUCs than planned dose-widths at 30-60 Gy, and the planned 65 Gy dose-width produced a greater AUC than the accumulated 65 Gy dose-width. Accumulated EUD was also predictive of G2 GI toxicity, where planned EUD was not.

Stool frequency (n = 34, 24.5%)

The accumulated 45 Gy dose-width was most strongly associated with increased stool frequency, with AUC = 0.648 [0.536, 0.759]. Planned dose-widths from 30 to 50 Gy were predictive of stool frequency, as were accumulated dose-widths from 30-60 Gy. Planned dose-widths from 30 - 40 Gy produced greater AUCs than corresponding accumulated dose-widths. For all other dosimetric parameters, accumulated AUCs were greater than planned AUCs.

Bowel bother $\geq$ G1 (n = 55, 39.6%)

The strongest predictor of bowel bother G1 was the accumulated 60 Gy dose-width (AUC 0.638 [0.543, 0.733]). Also predictive were planned dose-widths 40-50 Gy, accumulated dose-widths 50-65 Gy, and accumulated EUD. This is in contrast to discovery cohort analysis, which found no associations with bowel bother $\geq$ G1.

Bowel bother $\geq$ G2 (n = 32, 23.0%)

The only predictor of bowel bother G2 was the accumulated EUD (AUC 0.648 [0.539, 0.757]). This was also in contrast to discovery cohort analysis, which found no associations with bowel bother $\geq$ G2.

Bowel bother $\geq$ G3 (n = 13, 9.4%)

The only predictor of bowel bother G3 was the accumulated EUD (AUC 0.686 [0.533, 0.839]), which was greater than the accumulated EUD AUC from both G1 and G2 bowel bother. As severity of bowel bother increased, AUC also increased. However, width of 95% CIs also increased.

4.4 Discussion

DSMs of the rectal wall, calculated for planned and accumulated delivered dose, were parameterised using EUD and dose-widths for 139 patients in the consolidation cohort of the VoxTox research study. Previous results for 109 patients from the discovery cohort demonstrated for the first time that accumulated dose can be a better predictor of rectal toxicity than planned dose in patients receiving radiotherapy for prostate cancer. Toxicity endpoints reported in the analysis of the discovery cohort were selected based upon availability of baseline information. For patients in the consolidation cohort, formalised reporting systems were introduced into the VoxTox study, so prospective baseline data were available for all patients. Additionally, imaging and contouring were more standardised and robust due to the implementation and dissemination of study protocols and guidelines. A further difference between the analysis presented for the consolidation cohort here, and the discovery cohort in Section 4.2 was that the software for autosegmenation of the rectum on MVCT images was upgraded from version 1.4 to 1.6. Improvements were made in identifying the rectal contour in regions where air pockets were detected on the MVCT scan. Details of version 1.6 of the autosegmentation software have been presented in the literature [143].

Twelve endpoints were explored in the analysis of the consolidation cohort. Endpoints were selected based on the most commonly reported toxicities in the literature, as well as prevalence within the dataset. Increasing grades of severity (i.e. $\geq$ grade 1, $\geq$ grade 2) were included if the level of toxicity was experienced by greater than 5 % of patients.

Significant dose-toxicity associations were observed for 11 of 12 endpoints. The dose parameter most discriminative of toxicity was from accumulated dose, rather than planned

dose, for 8 of 11 endpoints. Spatial dose-widths (at various dose levels) were more predictive of toxicity for 7 endpoints, and EUD was found to be the strongest predictor for 4 endpoints.

Accumulated dose was lower than planned dose, when comparing EUD, for 90 % of patients. The previous discovery cohort analysis found all patients to have lower accumulated dose than planned dose. This result supports the suggestion that interfraction motion leads to a dose-blurring effects at higher dose levels when accumulating daily delivered dose. As with the discovery cohort, some 0 % dose-width results were recorded at higher dose levels. For planned dose, 2/139 patients recorded null values at the 70 Gy dose-width, and for accumulated dose, 2/139 and 7/139 patients recorded null values at 65 Gy and 70 Gy dose-widths, respectively.

Overall, the analysis suggests that stronger dose-toxicity associations are determined using accumulated dose, rather than planned dose. In general, spatial dose-widths produced stronger associations than EUD.

Where significant associations were observed from the discovery cohort analysis, either the same result was found (accumulated EUD was most associated with CTCAE $\geq$ Grade 1 rectal bleeding for both discovery and consolidation cohorts), or results were complementary (for grade 2 proctitis, the discovery cohort analysis found accumulated EUD to be the sole predictor, whereas the consolidation cohort analysis considered a different scoring system and found accumulated 60 Gy dose-width to produce the greatest AUC, but accumulated EUD was also significant).

Analysis from two separate cohorts of the VoxTox study, discovery and consolidation, support the hypothesis that accumulated delivered dose is more predictive of rectal toxicity in prostate radiotherapy than planned dose. However, the benefit of incorporating spatial dose information (dose-widths) over a one-dimensional dose descriptor (EUD) is not distinctly evident. Further investigation of dose parameterisation approaches more suited to interpretation of accumulated dose are recommended. By improving the understanding of accumulated dose distribution to the rectal wall, we aim to develop better models for predicting rectal toxicity.

4.5 Concluding discussion

Dose-toxicity analysis was performed for two separate cohorts of the VoxTox research programme. Dose surface maps of the rectal wall have shown that accumulated dose can be a better predictor of rectal toxicity in prostate radiotherapy than planned dose. The ability to preserve and accumulate spatial dose information throughout treatment is a novel process requiring careful consideration of data interpretation and parameterisation. Future work may involve exploring alternative methods for geometrical quantification of spatial dose distributions in order to determine stronger correlations with toxicity. Further analysis of delivered dose to the rectal wall could facilitate the identification of inhomogeneous intra-organ radiosensitivities, allowing shape-based dose constraints to be derived. Additionally, spatial considerations could complement current DVH-based approaches to treatment planning.

Through novel characterisation of delivered dose, beyond the limitations of the static planned DVH, the aim was to determine those parameters strongly associated with rectal toxicity which could be incorporated into multivariate NTCP models. Emerging dose quantifiers could be integrated into planning constraints, as well as being prospectively monitored throughout treatment. Delivered dose can be accumulated in 'real-time' and analysed with each fraction, allowing on-treatment toxicity risk assessment. Towards the end of the course of treatment, if toxicity prediction was found to be lower than planned, the decision could be made to increase the total delivered dose to the target. The potential scope for further individualisation and adaptation of treatment could ultimately reduce rates of toxicity incidence and improve clinical outcomes.

Chapter 5

Voxel-level dose analysis for interfraction motion of the rectum

In this chapter, biomechanical finite element (FE) analysis is introduced as a method for improving the accuracy of calculation and accumulation of dose to the rectum for prostate radiotherapy patients in the VoxTox study. In the previous chapter, dose surface maps (DSMs) were generated where analysis of rectal motion was limited to isotropic deformation in the plane of the 2D imaging slice. These will henceforth be referred to as 2D-DSMs, to distinguish from DSMs generated from biomechanical FE modelling, which will be referred to as FE-DSMs. Biomechanical FE modelling provides a more anatomically plausible solution for rectal motion by allowing anisotropic deformation in 3D. The material elements used to construct the model facilitate increased resolution of dose-history tracking, enabling voxelwise analysis of dose. This chapter is divided into three sections:

- *Biomechanical modelling for voxelwise dose analysis*, where details of generating the model based on VoxTox patient data are presented.
- *Quantifying interfraction motion of the rectum*, where the range of interfraction rectal deformation is evaluated across the full course of treatment. Geometric correlations and dosimetric associations are investigated.
- *Sensitivity of dose response with simulated rectal motion*, where discrete intervals of rectal expansion and contraction, as well as extremes of rectal deformation as measured from patient data, are simulated. The dose response is determined for two different clinical scenarios. Resulting implications for normal tissue complication probability (NTCP) are also presented.

The research questions being addressed in this chapter, through the use of biomechanical FE modelling, are: how does rectal positioning change during the course of radiotherapy, and how does rectal motion correlate with delivered dose, and hence NTCP?

The original biomechanical FE model and corresponding MATLAB code was developed by Professor Michael Sutcliffe, Cambridge University Engineering Department, and was previously adapted by former Masters students for experimental purposes. Within the scope of this PhD project, the code has been adapted to read and process patient DICOM files, model parameters have been optimised for the rectal wall, and the system has been integrated into the VoxTox workflow. Interfraction motion has been analysed based on the outcome of the autosegmentation system developed by Prof Sutcliffe (as described previously in Section 3.1.3). Systems for extracting and processing properties of contour data, and the design and analysis of the sensitivity study, now possible through the application of FEM, have all been developed as part of of this PhD thesis.

5.1 Biomechanical modelling for voxelwise dose analysis

Biomechanical models are physical-based simulations that incorporate material properties of anatomy in addition to geometry and boundary conditions [21]. Biomechanical finite element (FE) analysis enables tracking of soft tissues undergoing large deformations, and has previously been applied to rectal simulation in the context of investigating motion of the prostate [18]. An advantage of biomechanical FE modelling is that accuracy remains constant in low contrast imaging (such as MVCT), provided that the boundary conditions can be defined [21, 23, 162]. The FE models described here were based upon manually segmented rectal contours from the kVCT planning scan, and autosegmented rectal contours from the MVCT image guidance scans. Autosegmentation of image guidance scans was essential to this work, and has been identified as a crucial component for bringing FE modelling into the clinical workflow [21]. Model performance is limited by the accuracy of the segmentation of the anatomy that defines the boundary conditions applied to the model (see Section 3.1.3). The rectal model consists of material elements, acting as surrogates for 3D volumetric pixels, or voxels, and hence clusters of cells. By tracking the interfraction movement of each voxel, the accuracy and resolution of dose accumulation is improved with respect to the methods previously adopted in Chapter 4 due to the capability of calculating individual dose histories. The previous method of dose accumulation using two-dimensional dose surface maps (2D-DSMs) allowed only for isotropic in-plane expansion, where the rectal circumference was uniformly normalised. Biomechanical FE modelling provides a more anatomically plausible solution for tracking rectal deformation, and hence a more accurate

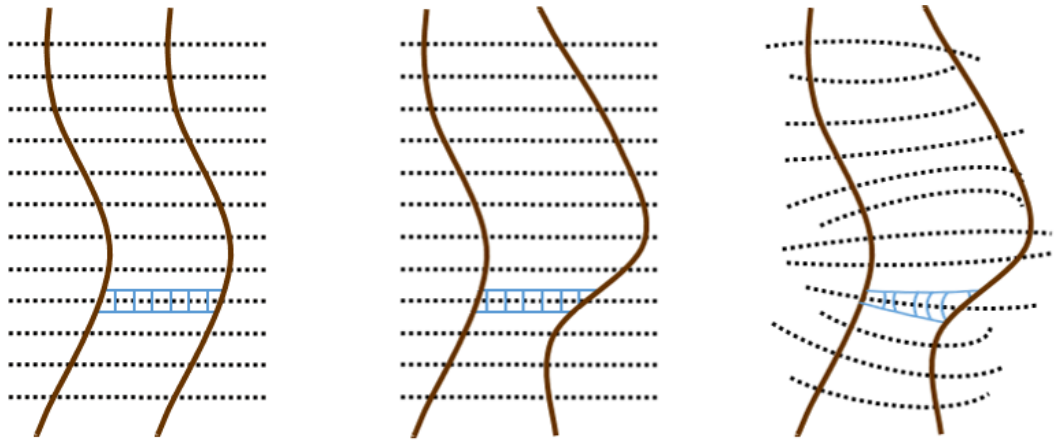


Fig. 5.1 Comparison of 2D and FE based approaches for generating DSMs. The rectal contour is shown in brown, black dotted lines correspond to rows in the resulting DSMs, and blue boxes correspond to voxels. (a) shows an example of the rectum from the reference kVCT treatment planning scan. For the 2D-DSM, voxels are normalised around the circumference of the rectum, in the plane of the image slice; (b) shows the deformed rectum at the time of treatment, and to generate the 2D-DSM, voxels are sampled in the same imaging plane, and uniformly normalised around the new rectal circumference in that plane; (c) demonstrates how FE-DSMs are generated, through out-of-plane expansion, and non-uniformly distributed voxels that deform according to material properties and applied boundary conditions. This presents a more anatomically representative model.

platform for delivered dose calculation [8]. This is illustrated in Figure 5.1. FE models of the rectum were generated in the Abaqus Computer Aided Engineering (CAE) Finite Element Analysis (FEA) environment (Dassault Systèmes). The process of generating and running the model for the VoxTox patient dataset is described within the following section.

5.1.1 Reading patient DICOM data

Rectal contours, whether manually contoured on kVCT images or autosegmented from MVCT scans, were stored as DICOM structure files. DICOM structure files must be read alongside the corresponding image set to be assigned a coordinate system. For each patient, the DICOM structure files of manually contoured and autosegmented rectal contours were imported into MATLAB alongside the respective kVCT and MVCT DICOM image sets for 20 or 37 fractions. Raw pixel coordinates were translated into millimetres by scaling by the image resolution (typically 272×272 pixels per slice, pixel size 1.953 mm for the downsampled kVCT; 512×512 pixels per slice, pixel size 0.754 mm for MVCT) and slice thickness (3 mm for kVCT and 6 mm for MVCT). When importing the MVCT rectal contours, an additional step was required in order to register the coordinate systems of the daily MVCT image set with the kVCT. The translational shifts applied during patient positioning at the

time of treatment were extracted and stored as private tags within the DICOM header; these made it possible to rigidly register the image sets. Any rotational offsets applied at time of treatment by small adjustments to gantry rotation were accounted for within the dose calculation step.

5.1.2 Model geometry

The simulated rectal model was constructed of 'innerpart' and 'outerpart' structures, as indicated in Figure 5.2. The outerpart represents the rectal wall, and the innerpart is the structure to which a displacement loading is applied. The innerpart and outerpart were constructed of material elements (voxels) and nodes (element vertices). The innerpart used surface membrane elements with 4 vertex nodes, negligible thickness, and reduced integration (element type SFM3D4R), with 30 nodes around the circumference of the rectum, and 120 nodes along the length (the patient's cranio-caudal z-direction). The outerpart, or rectal wall, was constructed of stress/displacement elements (type C3D8); fully integrated 8-node linear brick of initial thickness 3 mm and material density 0.05 g/cm^3 [18, 36]. The outerpart was assigned 80 nodes along the rectal length, and 30 nodes around the circumference. A previous study demonstrated no advantage to increased model resolution; increasing the number of elements by a factor of 36 not only increased computational time from 1 to 34 minutes, but had no significant effect on model accuracy [24]. Here, the resolution of the outerpart was assigned to approximate the number of voxels around the circumference as used in the 2D-DSM study (Chapter 4), but with increased resolution along the length of the rectum in order to accurately track out-of-plane motion along the rectal length (or z-direction). This was previously not possible to account for using the 2D-DSM approach.

The starting geometry of the innerpart was based on a 'central mandrel' - a cylindrical structure of radius 10 mm. This was common to all patients, and provided the basis for intra and inter-registration between fractions and patient contours allowing tracking of individual voxels. The coordinates of the patient's kVCT rectal contour were used to assign coordinates in space to the central mandrel. The centre of mass of each contour was calculated and then smoothed; the resulting coordinates formed a 'centreline'. The cylindrical contour of the central mandrel was then centred about this centreline. The centreline-adjusted central mandrel formed the starting geometry for each patient's FE rectal model and was used as a starting point for the kVCT based model, as well as each subsequent daily MVCT model generated for that particular patient, as shown in Figure 5.2.

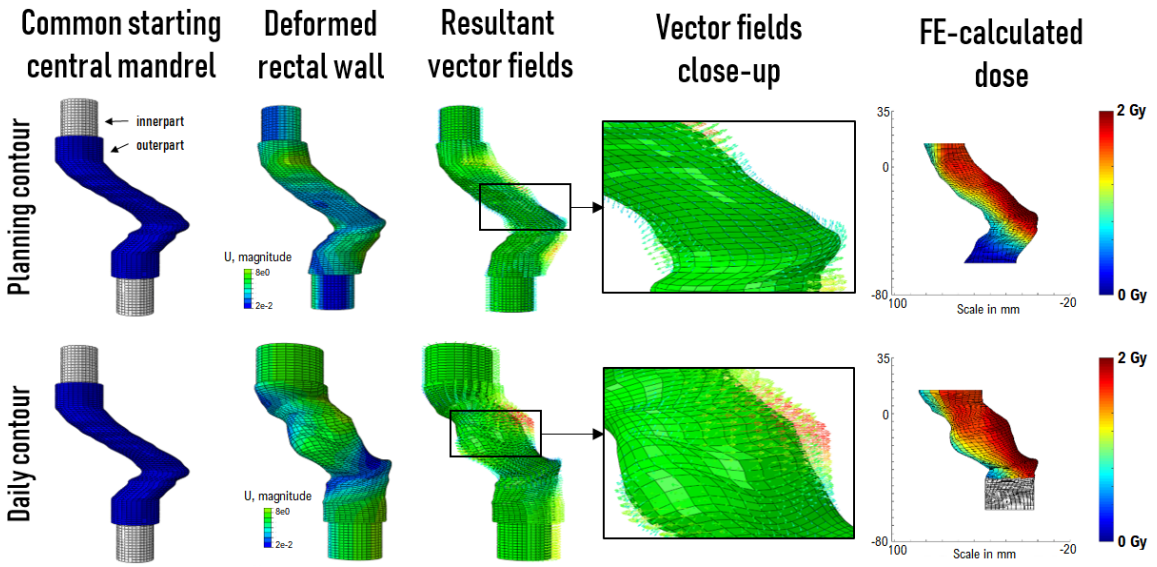


Fig. 5.2 Finite element (FE) model generation for the kVCT-based planning contour and a MVCT-based daily contour for an example patient. The patient's centreline-adjusted central mandrel shows the original 10 mm radius cylindrical central mandrel offset by the centreline of the kVCT planning contour, which provides the starting geometry for all models generated for this patient. The deformed rectal wall shows the final output of the biomechanical FE model. The corresponding resultant vector fields illustrate the out-of-plane displacement, which could not be modelled by the previous 2D-DSM approach in Chapter 4. FE-calculated dose shows the planned and daily delivered dose based on the anatomical changes of the rectum. The effect of rectal motion on dose is addressed in this chapter. Differences between planned and accumulated delivered FE-DSMs are investigated in Chapter 6.

5.1.3 Hyperelastic modelling

Hooke's Law can be used to solve simple linear elastic problems with material constants such as Poisson's ratio (ratio of lateral strain to longitudinal strain) and Young's modulus (elasticity). However, for anatomical modelling, soft tissues demonstrate a non-linear elastic stress-strain relationship [65] which can be better represented using hyperelastic constitutive laws. These depend on a strain-energy function (U) describing the energy stored in the material caused by deformation [21]. Here, the rectal wall was assumed to be an incompressible rubber-like material and was modelled using an isotropic hyperelastic model. The neo-Hookean model was applied, with initial model parameters normalised such that the shear modulus μ_0 (measure of rigidity) and bulk modulus, K_0 (measure of compressibility) corresponded to a Poisson's ratio of 0.49 [18]. The shape of the stress-strain curve, rather than the absolute values of the applied material properties, determines the mechanical behaviour of the rectal wall as it deforms. Many hyperelastic models have been adopted for modelling of rubberlike materials in FE simulations in the literature [68]. Of the physical-based models, the popular choices for anatomical modelling have been: the Mooney-Rivlin

model, reported to be of high accuracy and efficiency; and the Ogden model, an efficient and well-established model but dependent on 6 experimentally-derived material parameters [68]. The neo-Hookean model mimics the Mooney-Rivlin model with a single material parameter, and was selected here for its simplicity and applicability over a strain range appropriate for rectal expansion.

5.1.4 Loading and boundary conditions

For the VoxTox rectal simulation, the focus was on strains and deformations rather than stresses. Therefore, displacement loading was applied which expands the model based on geometrical information from the rectal contour data, rather than by applying a force. This approach is recommended due to the difficulties in measuring applied forces [21]. The accuracy of the defined boundary conditions has a greater impact on model performance than the applied elastic modulus [21, 134]. Final FEA solutions are computed per node based on the material properties and boundary conditions.

Boundary conditions are defined from the coordinates of the manual or autosegmented rectal contours. Nodal coordinates were determined by converting DICOM rectal contour coordinates to a polar system and using spline interpolation. Polar coordinates were translated back to the Cartesian system. The displacement loading was the difference between the final contour position and the corresponding nodal coordinate of the centreline-adjusted central mandrel, and was applied as a uniformly increasing load (Figure 5.2). The contact interaction between innerpart and outerpart was defined as a frictionless sliding surface. Further boundary conditions were applied at the most superior and inferior extremes of the rectal contour, which allowed the model to move in the patient's cranio-caudal z-direction by ± 3 mm, accounting for the uncertainty in defining these edges of the rectum inherent to the MVCT slice thickness. This also corresponded to the rectal displacement measured in the cranio-caudal direction by Boubaker et al. [19].

5.1.5 Running the model in Abaqus

Model parameters were processed and combined in MATLAB to generate an input file for FE modelling in Abaqus. The input file used the above information to define:

- Geometry (node numbers and coordinates, element type, surface density, for each part)
- Assembly (initiating parts and instances)
- Material (apply hyperelastic model properties, density)

- Boundary conditions (defining range of motion restrictions)
- Interaction properties (i.e. define sliding surface behaviour, set displacements and loading amplitude)
- Steps (i.e. non-linear, dynamic, explicit)

The final simulated rectal model was deformed according to these properties. Batch processing was performed using DOS commands. The coordinates of the final deformed outer surface of the rectal wall were automatically extracted using a Python script.

5.1.6 Post-processing and dose calculation

The patient's DICOM radiation dose file corresponding to the day of imaging was imported into MATLAB. For daily dose, the treatment couch shifts were applied to register to the kVCT coordinate system (as with the MVCT images and structures during pre-processing of model geometry). Any corrections for roll during treatment (implemented by gantry rotation), were also incorporated into the dose matrix. Dose was interpolated in 3D at each of the 2400 final node positions on the surface of the outerpart and displayed on the 3D contour (as shown in Figure 5.3). Similar to the 2D-DSM approach, where data at the superior and inferior extremes of the rectum were cropped due to the limited MVCT field of view, missing data were substituted with planned dose during daily dose accumulation. By starting from a common starting structure, each element was tracked through its daily deformation, allowing a greater level of accuracy and resolution in the calculation and accumulation of delivered dose than previously achievable.

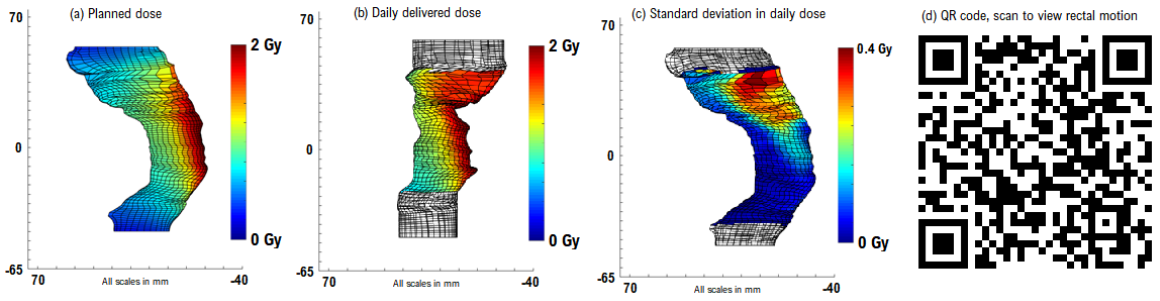


Fig. 5.3 Biomechanical FE model showing rectal motion; (a) Planned dose based on kVCT contours, (b) Daily delivered dose calculated on MVCT contour, (c) Standard deviation across 37 fractions shown on kVCT planning contour, (d) QR code link to rectal motion example. QR code can be read using the camera application on a tablet or smartphone. Focus camera on code. QR reader may need to be installed.

5.2 Quantifying interfraction motion of the rectum

5.2.1 Introduction

The hypothesis of the VoxTox research programme is that delivered dose to the rectum is more predictive of rectal toxicity in prostate radiotherapy than planned dose. Part of the motivation for this comes from studies reporting significant rectal motion in patients undergoing prostate radiotherapy [43, 117]. When the rectal position at time of treatment differs to that identified on the pre-treatment planning scan, this introduces deviations between the intended planned dose to the rectum as accepted at the time of treatment planning, and the dose actually received by the rectum.

Daily interfraction motion of the rectum can be visualised on daily image guidance scans, but to date has been difficult to quantify due to the quality of the images. In particular, segmentation of anatomy within the pelvic region can be challenging due to soft tissue boundaries which lack distinction and worsen in low contrast imaging [67, 90]. Methods used for segmentation of the prostate, such as deformable image registration (DIR), are generally not applicable for the rectum due to the large and unpredictable spatial deformations caused by rectal contents and intestinal gas [97, 114, 136, 161]. Common DIR algorithms struggle due to intensity variations and the lack of one-to-one correspondence between a full, gassy, or empty rectum [40, 114, 175]. Previous studies investigating the dosimetric effects of interfraction rectal motion have been dependent upon manual delineation of the rectum on image guidance scans [11, 43, 75, 85, 96, 125, 145], and are consequently limited in sample size.

Within the VoxTox research programme, the autosegmentation software was developed by Professor Michael Sutcliffe, Cambridge University Engineering Department. An automated solution for identifying the rectum on daily image guidance scans was vital for the contouring of over 56000 slices. A full description of the autosegmentation algorithm was published by Shelley & Sutcliffe in Biomedical Physics & Engineering Express [143].

The following section presents results of quantitative analysis of the magnitudes of rectal motion throughout the course of 37 fractions of prostate radiotherapy as defined using the VoxTox autosegmentation system. The geometric measures of rectal deformation are explored for associations with delivered dose as calculated and accumulated using finite element dose surface maps.

5.2.2 Material and methods

Seventy-six patients treated with 37 fractions were selected from the VoxTox consolidation cohort. Patients were prescribed 74 Gy to the prostate, with daily image guided radiotherapy (IGRT) delivered using TomoTherapy®. Baseline rectal measurements were extracted from the kilovoltage computed tomography (kVCT) planning scans, which were manually contoured by experienced clinical oncologists according to the VoxTox protocol. Imaging specifications for the downsampled kVCT were: 272×272 pixels per slice, pixel size 1.953 mm, slice thickness 3 mm. Scan length included the full extent of the rectum. Rectal length was defined from the rectosigmoid junction, at the last slice where the bowel turns anteriorly and to the left, to the most inferior slice containing both ischial tuberosities [136]. Images were inspected in both the axial and sagittal planes. Measured rectal length was taken as the number of kVCT slices containing rectal contours multiplied by the slice thickness. Daily rectal contours were determined from the megavoltage computed tomography (MVCT) image guidance scans using in-house autosegmentation software [143] based on the Chan-Vese algorithm. MVCT specifications were: 512 × 512 pixels per slice, pixel size 0.754 mm, slice thickness 6 mm. The field of view for MVCT imaging was limited to typically 8–12 slices according to local protocols to minimise additional dose and time for prostate IGRT [15], so only a proportion of the rectum was imaged.

For each kVCT or MVCT slice containing a rectal contour, the equivalent diameter of the contour was extracted using MATLAB's *regionprops* function. The equivalent diameter is defined as the diameter of a circle with the same area as a binary mask fitted to the contour, calculated as $\sqrt{(4 \times Area/\pi)}$.

Due to the use of daily image guidance, the prostate was assumed to be a fixed landmark. A reference plane containing the most anterior point of the prostate contour as identified on the planning scan, was used to define a surrogate z-origin. This approach was recommended by Dr Rajesh Jena, consultant clinical oncologist within the VoxTox research programme, as opposed to using mid-prostate, for example, which depends upon the accuracy of identifying the superior and inferior boundaries of the prostate. Defining the origin in this way presented a solution for registering contour data from daily MVCT images with the planning kVCT scan, and across multiple patient datasets.

Daily MVCT image guidance scans were corrected for couch positional adjustments applied at the time of treatment. This information was stored in private DICOM tags and used to apply rigid transformations in x, y, and z-directions [132]. Once registered to the frame of reference of the planning scan, the MVCT slice closest to the reference plane, was defined as the z-origin. Due to differences in slice thickness (kVCT 3mm v MVCT 6mm), it was possible that the MVCT z-origin may differ by up to 3mm with respect to the kVCT z-origin,

and for a discrepancy of up to 6 mm to exist between the z-origins defined on MVCT images acquired on different days.

Due to differences in daily MVCT scan lengths and positioning, less data were available as slice position from the z-origin increases. Contour measurements were acquired at each slice location, provided that there were data from more than 15 patients (approximately 20 %), and mean measurements per patient were calculated from a minimum of 5 fractions. This restricted analysis to ± 7 slices about the z-origin. A positive slice number indicates the cranial direction. Where applicable, mean values, standard deviations, and ranges of motion were recorded for each patient across the 37 fractions. The standard deviation across all fractions was taken as a measure of rectal motion for each patient.

Associations between continuous variables were assessed using the Pearson correlation coefficient, r, where 0 indicated no relationship, and +/-1 indicates a perfect positive/negative linear relationship. Correlation coefficients between 0.2-0.39 were considered weakly correlated, and between 0.4-0.59 were considered moderately correlated. The significance of the correlation was indicated by the accompanying p-value test statistic.

5.2.3 Results

5.2.3.1 Geometric correlations

The mean rectal length measured from the kVCT planning scan was 98.9 mm (SD 16.3, range [66, 165] mm). This did not incorporate the systematic error of +/- 1.5 mm at superior and inferior boundaries, or the effect of multi-observer contouring (discussed by Scaife et al. in [138]). Results were similar to Marcello et al. [92], who found the median rectal length for 709 patients to be 95 mm. Rectal length was plotted against height for all patients as shown in Figure 5.4. No significant correlation was observed between height and rectal length (Pearson correlation coefficient r = -0.009, p = 0.939).

Overall, rectal length was not found to be correlated with equivalent diameter from kVCT, or MVCT mean or maximum equivalent diameter, suggesting that longer rectums are not prone to larger lateral expansion. However, from MVCT slices -1 to +7, weak to moderate negative correlations were observed between rectal length and minimum MVCT equivalent diameter, suggesting that longer rectums tend to narrow towards the superior portion of the rectum. Rectal length was moderately correlated with rectal motion (quantified by the standard deviation of the MVCT equivalent diameter throughout treatment) from slices -1 to -4, i.e. in the middle to inferior section of the rectum, suggesting that longer rectums tend to move more in the lower aspect.

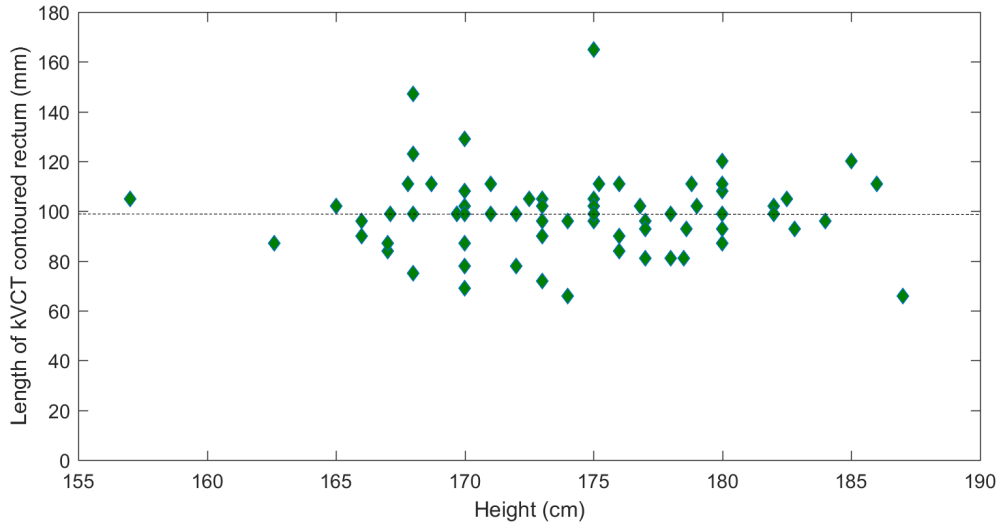


Fig. 5.4 Plot of patient height versus rectal length as defined on the manually segmented kVCTs scan for 76 patients. Dotted line indicates line of best fit, m = 0.026.

Differences between equivalent diameters calculated from kVCT contours and mean, minimum, and maximum MVCT contours averaged across all fractions are shown in Figure 5.5. Quoted differences between kVCT and MVCT contours are quantified using differences in equivalent diameter. The kVCT contours were significantly different to the corresponding mean MVCT contours at central slices from slices -2 to +4, and at the most superior and most inferior slices, +7 and -7 respectively (paired samples t-test, $p<0.05$). Minimum MVCT contours were significantly different from corresponding kVCT contours for slices -5 to +7. Maximum MVCT contours were significantly different from corresponding kVCT contours across all slices. Additionally, both minimum and maximum MVCT contours were significantly different to the mean MVCT contour across all slices. This suggests that the change in contour size due to interfraction motion throughout the course of treatment can be considered a significant effect. No correlation was found between the equivalent diameter of the planning kVCT contour and rectal motion (standard deviation of MVCT contour equivalent diameters), at any slice. This suggests that the rectal contour used for treatment planning is not indicative of the subsequent change in rectal contour due to interfraction motion throughout the course of treatment, despite the use of image guidance (i.e. a narrower rectum at treatment planning will undergo the same amount of interfraction motion as a larger rectum throughout the course of treatment).

The rectum is often described as being divided into thirds; superior (sup), middle (mid), and inferior (inf) [136]. Anatomically, the superior third connects to the sigmoid, the middle third is generally at the level of the prostate, and the inferior third connects to the anal canal.

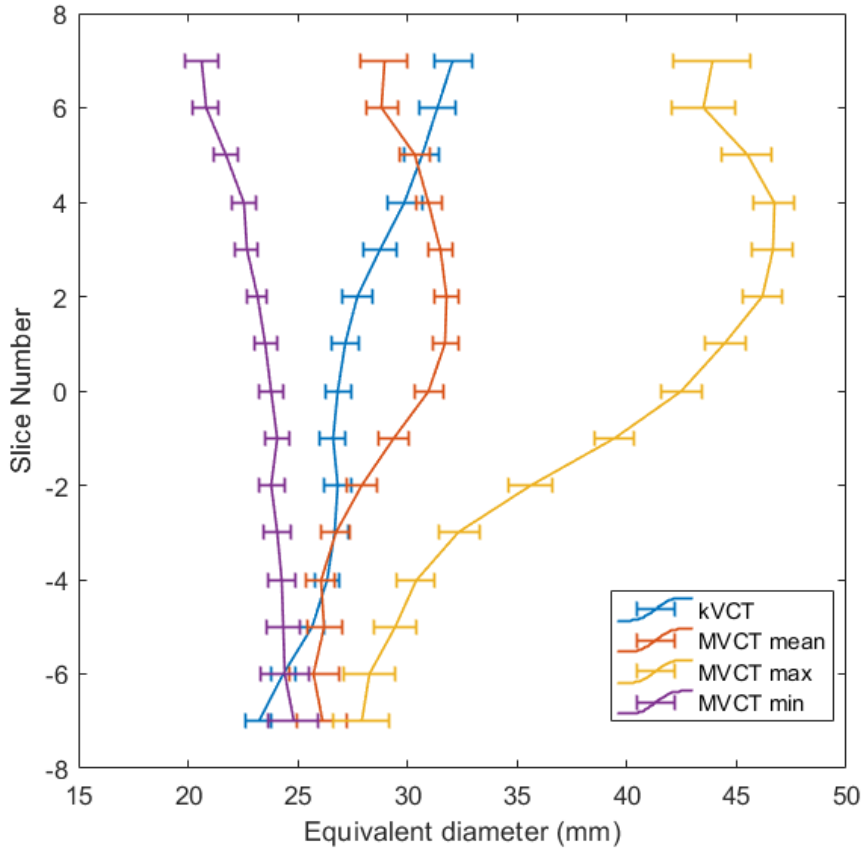


Fig. 5.5 Change in rectal equivalent diameter throughout the course of treatment. kVCT data are based on manual segmentation, MVCT data are based on autosegmentation. Equivalent diameter is calculated at each slice where sufficient data were available. The 0 slice is defined as the registration origin; positive slice number indicates cranial direction, negative slice number indicates caudal direction. Slice thickness is 6 mm. MVCT mean, max, and min equivalent diameters are averaged across all patients and error bars represent the standard deviation in the overall average. Slices +6, 0, -6 were used as surrogates for the superior, middle, and inferior rectal thirds, respectively.

To investigate whether the rectal thirds behave differently, rectal geometries of sup and inf thirds were compared with the mid-third. For a mean rectal length of approximately 10 cm, contours from the MVCT +6 slice (+36 mm) were used as surrogates for the sup-third, the zero slice for the mid-third, and the -6 slice (-36 mm) for the inf-third, shown in Figure 5.5.

For kVCT contours, equivalent diameter was significantly narrower at the inferior third (-2.6 mm [SD 4.6], $p<0.001$), and wider at the superior third (4.5 mm [SD 5.8], $p<0.001$), with respect to the zero slice mid-third. However, from MVCT contours, the mean equivalent diameter measured across all patients and all fractions, were narrower at the inf and sup thirds than the mid-third. Mean difference at the inf-third with respect to the mid-third was -5.5 mm (SD 7.2, $p<0.001$), but the difference between sup and mid-third was not significant. For

minimum equivalent diameter, the difference between inf and mid third was not significant, but the sup third was significantly narrower than the mid third by 2.6 mm (SD 5.4, p = 0.001). The maximum equivalent diameter was significantly narrower at the inf third than mid third (-15.5 mm [SD 11.2], $p \leq 0.001$), but differences were not significant between sup and mid thirds.

Rectal motion was significantly less at the inf third than mid third. The standard deviation of the measured equivalent diameters across all fractions and patients was 4.0 mm lower at the inf-third than mid-third ($p \leq 0.001$). In contrast, rectal motion at the sup-third was greater than at the mid-third, with a difference in standard deviation of 1.5 mm ($p \leq 0.001$). The standard deviation of equivalent diameters measured across all fractions and patients increased towards the superior/cranial direction. Results suggest that the sup-third demonstrates greater rectal contour deformation due to interfraction motion than the mid-third, and that the inf-third undergoes less interfraction motion than the mid-third which may be due to the stiffness of surrounding musculature. Therefore, substituting planned dose in areas where daily delivered dose information is missing due to field of view limitations may be more appropriate at the inferior end of the rectum than at the superior end. However, this conclusion is extrapolated beyond the data presented here.

5.2.3.2 Dosimetric associations

Rectal length was weakly negatively correlated with both planned EUD (r = -0.278, p = 0.02) and accumulated EUD (r = -0.308, p = 0.01), suggesting that longer rectums receive a greater proportion of low dose contribution leading to a lower overall EUD. Correlations with rectal length may be weak due to the low variance in measured rectal length. Equivalent diameters from kVCT planning contours were not found to be associated with planned EUD. kVCT equivalent diameter at the 0 and -1 slice were weakly negatively correlated with the 50 Gy dose-width (r = -0.236, p = 0.042 and r = -0.240, p = 0.038, respectively). Overall, min/mean/max equivalent diameters from MVCT contours, and rectal motion over the course of treatment, were not correlated with accumulated EUD. Mean MVCT equivalent diameter at the -5 and -6 slice were weakly negatively correlated with the 50 Gy dose-width (r = -0.307, p = 0.041 and r = -0.397, p = 0.030, respectively), indicating that an intermediate dose level of 50 Gy may be affected by rectal expansion.

Rectal motion in the mid to sup portions of the rectum (slices +2 to +6) were weakly and negatively correlated with accumulated 60 Gy and 65 Gy dose-widths, supporting the suggestion that increased rectal motion leads to a blurring effect at higher dose levels.

The difference between kVCT and mean MVCT equivalent diameters at slices -2 to +5 were weakly to moderately correlated with the difference between planned and accumulated

EUD (range $r = 0.447$, $p < 0.0001$ at slice 0, to $r = 0.234$, $p = 0.047$ at slice +4). This indicates that in the mid to sup region of the rectum, an increase in rectal diameter is proportional to an increase in EUD, and therefore it may be possible to predict the change in dose based on the change in diameter.

5.2.4 Discussion

Using the automated segmentation system developed within the VoxTox programme, it was possible to extract geometric properties of the rectum from daily MVCT image guidance scans and present information on rectal motion that has not previously been extractable from daily image guidance scans. This enabled tracking of inter-fraction rectal motion through expansion and contraction, throughout the course of radiotherapy. This has been a challenging task in the past due to the poorer image quality of MVCT scans, and the sheer quantity of contours required (over 30,000 contours analysed within this sub-study). Significant differences were found between the geometry of the rectal contour accepted at the treatment planning stage, and the mean rectal position at the time of treatment delivery. The amount of rectal motion was evaluated within discrete thirds - sup, mid, inf - and it was found that the mid to sup thirds suffer more rectal motion than the most inferior third which was found to be narrower and more fixed by the surrounding musculature.

A change in rectal contour results in a change in the dose received by the rectum with respect to planned dose. Contour geometries were analysed for correlations with dose. Although the equivalent diameters of kVCT and MVCT contours were not correlated to planned and delivered EUD, respectively, both kVCT and MVCT results indicated that a narrower contour resulted in a larger 50 Gy dose-width. Greater rectal motion measured on MVCT contours resulted in a reduction of the 60 and 65 Gy dose widths. The difference in equivalent diameter from kVCT to MVCT contours is correlated with the difference between planned and delivered EUD.

Geometrical contour data was analysed to quantify the bounds of rectal motion, and was also considered in terms of the impact on delivered dose. Interfraction motion of the rectum was shown to be significant with respect to baseline planning data, inducing differences between planned and delivered dose. These results strengthen the motivation for monitoring interfraction motion to calculate delivered dose, which could ultimately act as a trigger for adaptive radiotherapy.

5.3 Sensitivity of dose response to simulated rectal motion

5.3.1 Introduction

The purpose of this study is to use physical-based biomechanical FE modelling to improve understanding of the effect of rectal deformation on dose, and how this translates to NTCP. In prostate radiotherapy, daily image guidance is used to confirm that the tumour is being accurately targeted, but little consideration is given to the rectum other than a visual inspection for excessive dilation. Due to variations in rectal content from day-to-day, the position, size, and shape of the rectum can differ from the original rectal contour delineated on the kVCT planning scan formulating the basis of the radiation treatment plan. This can result in deviations between the planned dose to the rectum accepted at the time of treatment planning, and the dose actually delivered (as demonstrated in Chapter 4). Quantifying the effect of variations in rectal contour on dose has been challenging to date.

In Section 5.2, the range of rectal deformations across the course of 37 daily treatment fractions was measured for a subset of VoxTox patients. Here, this information is applied to rectal simulations which deform baseline rectal contours within the measured range of motion. Simulations include: 'uniform' expansion and contraction, where the same expansion coefficient is applied to each slice of the rectal contour along the full length of the rectum; and 'non-uniform' expansion, where applied coefficients vary per slice depending on rectal properties measured from patient data. The former provides more information on motion-dose sensitivity, whereby a known shift is applied to all slices and the dose-effect can be measured at each increment, and the latter is more representative of the way the rectum behaves in reality based on measured patient data, i.e. rectal expansion is more limited in the inferior third than the superior third, for instance.

For each expansion/contraction simulation, two different dose scenarios are investigated. The first, scenario (i), assumes that the motion of the rectum affects the position of the prostate and, assuming this is corrected for via the use of image guidance, the dose is therefore shifted accordingly by the magnitude of simulated radial increase to the rectum. The second, scenario (ii), assumes that rectal motion does not affect the prostate position, i.e. the rectum expands into the high dose region causing the prostate to compress or deform, and so no dose shift is applied to the simulation. Both scenarios are clinically viable and are illustrated in Figure 5.6. For each simulation and dose scenario, the impact on NTCP was quantified based on the resulting equivalent uniform dose.

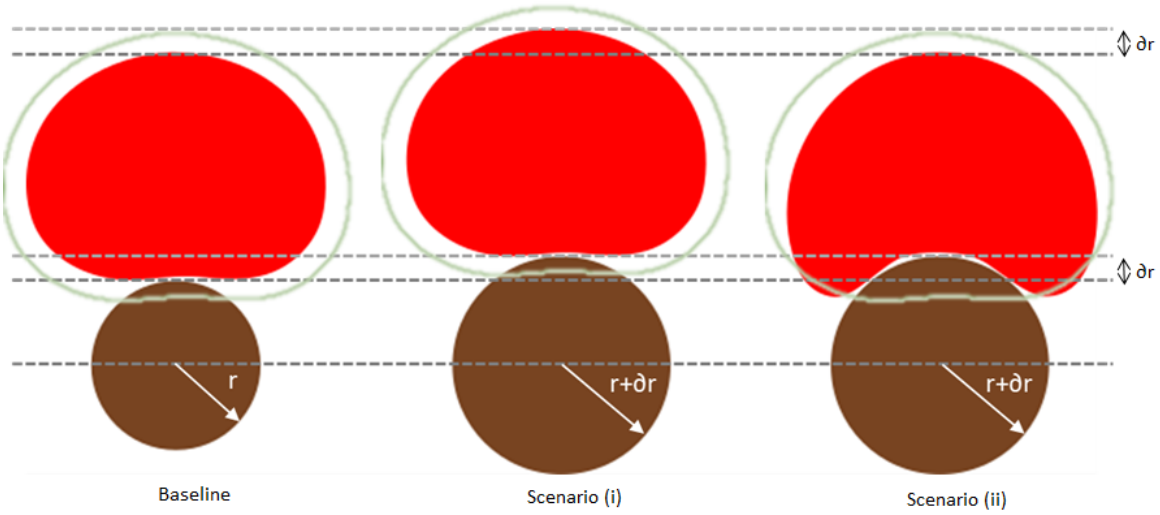


Fig. 5.6 Rectal (brown) expansion scenarios and impact on prostate (red) positioning, and subsequently dose (prescription isodose represented by light green line). With respect to baseline, scenario (i) assumes rectal expansion by δr pushes the prostate anteriorly by δr. It is assumed that the dose is correctively shifted following image guidance (i.e. the patient couch position would be dropped by δr). Scenario (ii) assumes the same rectal expansion, but rather than a rigid shift to the prostate, the rectal expansion causes the prostate to deform in shape, effectively squashing it so the edges become 'dog-eared'. The dose would then not be shifted with image guidance, and the rectum would have expanded into the high dose region.

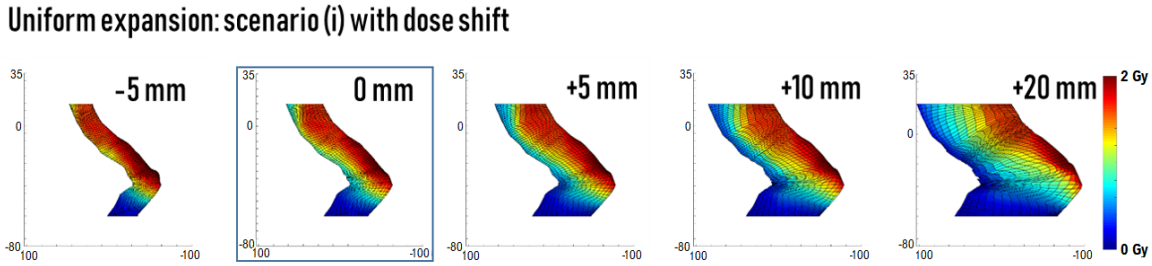
5.3.2 Materials and methods

Biomechanical rectal models were generated for 26 patients from the kVCT planning data, and the dose was calculated to the FE-DSM. These patients were a separate cohort to the patients investigated in the motion study in Section 5.2. Different rectal deformation scenarios were simulated in Abaqus, as detailed below. Deformations were applied to the DICOM RT STRUCT contours in MATLAB at the pre-processing stage whilst generating the input file for Abaqus. The model was constructed following the methods described in Section 5.1. Final simulation positions were exported so that resulting delivered dose could be calculated in MATLAB. Due to the different starting contours for each patient, dose differences were reported relative to the baseline dose, and were parameterised using EUD and maximum point dose.

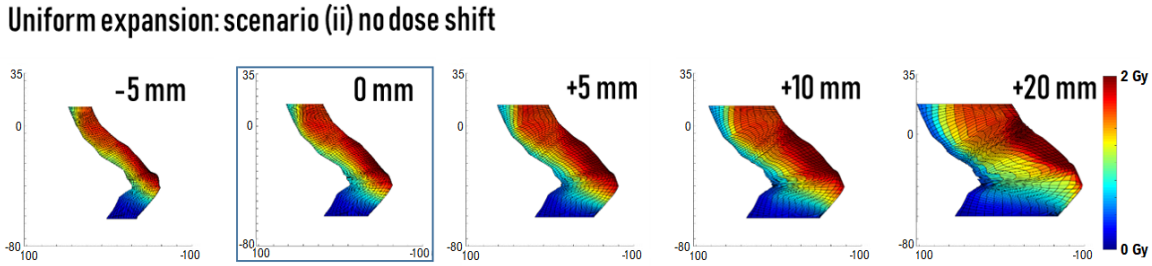
5.3.2.1 Uniform expansion

Rectal contours from the kVCT planning scan were expanded radially, by 1, 2, 3, 5, 8, 10, 15, and 20 mm, and contracted by -1, -2, -3, and -5 mm, uniformly across all planes of the rectal contour. This is illustrated for a sample of contraction and expansions for a single patient in Figure 5.7. Values were selected to incorporate a representative range of rectal motion.

The minimum diameter across all rectal contours for all patients was 13 mm, so simulating contractions of the rectal radius by an absolute value greater than 5 mm was not possible. For each patient, the coordinates of the rectal contour were extracted from DICOM structure files, and radially expanded in a polar coordinate system. Expansions were applied uniformly to each slice along the full length of the rectum. This resulted in 13 simulations per patient, including a baseline measurement with zero expansion.



(a) FEM simulation of uniform expansion where dose is shifted by expansion coefficient, scenario (i)



(b) FEM simulation of uniform expansion where no dose shift is applied, scenario (ii)

Fig. 5.7

5.3.2.2 Non-uniform expansion

For a separate patient cohort of 76 patients, as defined in Section 5.2, geometric measurements were taken from rectal contours at the reference slice (the slice containing the most anterior point of the kVCT prostate contour) and 7 slices in both the superior and inferior directions, on each of the 37 daily MVCT scans. The number and location of slices within each fraction's MVCT scans set varies, with a mean of 12 slices per scan [range: 8-17]. Data were included where slice measurements were available for 15 or more patients, each containing 5 or more scan sets (i.e. a combined minimum of 75 contours per slice). The purpose of performing simulations based on extremes of measured patient contours was to determine the maximum dose deviation from baseline for a single fraction, and to explore the approximate accumulated effect based on the mean measured contour position.

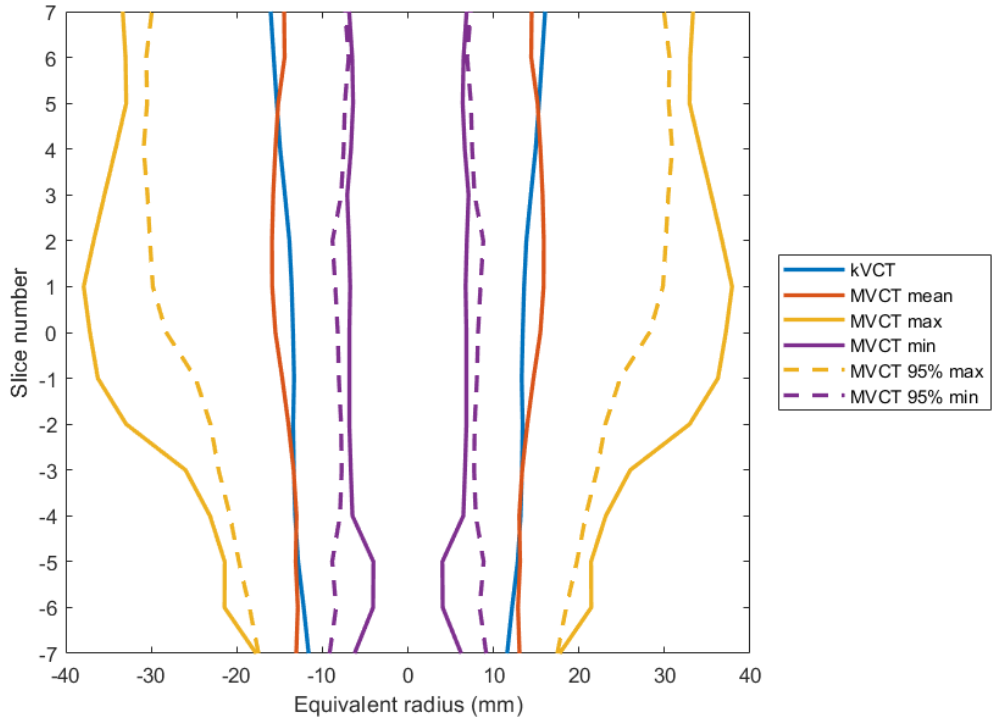
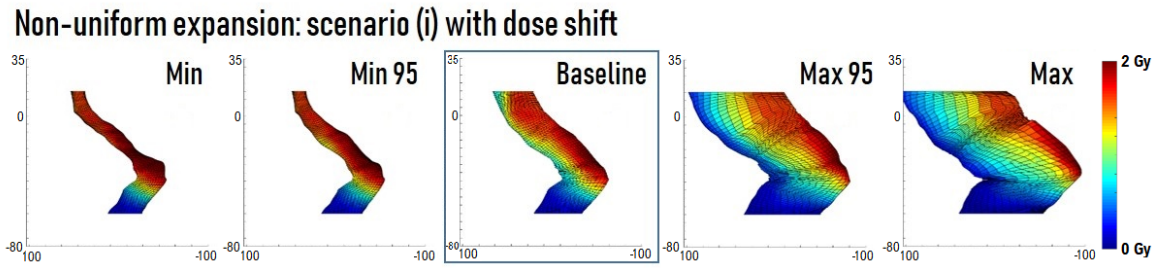
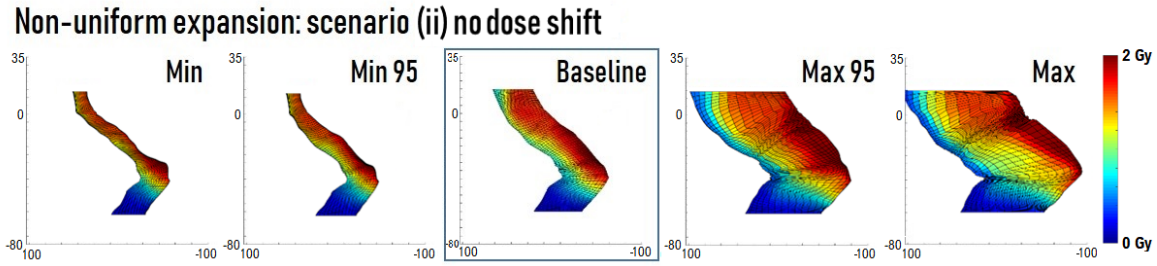


Fig. 5.8 Radial rectal expansions as measured from 76 patients, as described in Section 5.2. Radius dimensions are mirrored about the origin to mimic the physical cross section of the rectum. Mean radii were measured across all fractions and patients, for each slice of the rectal contour. Maximum and minimum values were the largest measured absolute expansion or contraction from baseline from any fraction or patient. The 95th percentiles (95% max and min) were calculated from the max and min across all patients to remove the effect of outliers. The zero slice is the registration origin, and positive slice number indicates cranial direction, negative slice indicates caudal direction. Slice thickness was 6 mm.

Using MATLAB's standard *regionprops* function, the equivalent diameter for each slice was calculated as the diameter of a circle with the same area as the segmented contour. Across all patients and fractions, the mean, global maximum, and global minimum equivalent diameters were measured per slice. Global maximum and minimum represented the most extreme measurements across all patients (with final slicewise values not necessarily from the same fraction or patient). To reduce the effect of outliers, the 95th percentiles of maximum and minimum values across all patients were also calculated for each slice. For each of the 5 sets of parameters (mean, maximum, 95th percentile maximum, minimum, 95th percentile minimum), the differences between MVCT and kVCT equivalent diameters were calculated and halved to find magnitudes of radial expansion to be applied to the rectal simulation. Radial expansions were non-uniformly applied along the length of the rectum, i.e different expansions were applied at each slice as shown in Figure 5.8, to the kVCT planning contours for 26 patients in this sensitivity analysis. An example is shown in Figure 5.9.



(a) FEM simulation of non-uniform expansion where dose is shifted by expansion coefficient at the zero slice, scenario (i)



(b) FEM simulation of non-uniform expansion where no dose shift is applied, scenario (ii)

Fig. 5.9

5.3.2.3 Dose shifting

Each set of simulations described above (uniform and non-uniform expansions) was performed twice in order to model different clinical scenarios in terms of dose delivery; scenario (i) the planned dose is shifted correspondingly, scenario (ii) the planned dose position remains static, unaffected by the rectal deformation. Because daily image guidance was used, target coverage was achieved through matching to the planning target volume (PTV) [81] contour via positional couch shifts or gantry rotations. Both scenarios are illustrated in Figure 5.6.

Scenario (i), assumes that rectal expansion induces prostate motion, i.e. when the rectum expands (or contracts), the prostate is consequently shifted anteriorly (or posteriorly) by the same magnitude. The planned dose would then be shifted and matched to the prostate, which in practice is achieved by altering the height of the treatment couch. This would be the case in the region where the rectum is connected to the prostate via a thin layer of fibrous tissue, the Denonvilliers' fascia [19]. In this investigation, this effect was achieved for uniform expansion by shifting the dose by the magnitude of the radial expansion, and for non-uniform expansion by shifting the dose by the magnitude of the radial expansion observed at the reference slice.

Scenario (ii) assumes that the rectum moves independently of the prostate, i.e. where the rectum is adjacent to the prostate, and the prostate position remains fixed, the rectum can

expand into 'prostate space', effectively compressing the prostate or causing the prostate gland to become 'dog-eared' in shape (see Figure 5.6). This would cause the rectum to shift further into the high-dose region, and no adjustment to the planned dose position would necessarily be made as target coverage is priority. Further from the prostate, this assumption allows rectal expansion without affecting the position of the prostate (or planned dose). In the case of contraction in scenario (ii), the assumption is that the rectum contracts away from the prostate without affecting its position. Both scenarios are clinically plausible and represent the extremes of a range in which the real clinical situation lies within.

5.3.2.4 Normal tissue complication probability

To understand the clinical implications of the simulated dose variations caused by rectal expansion and contraction, dose differences were translated into NTCP. A univariate binary logistic regression model was developed using EUD as the input variable, as a predictor of CTCAE Grade 2 rectal bleeding or above. The model was constructed following approaches by Schaake et al. [140], and parameters (Eq. 5.2) were optimised on data from 109 discovery cohort patients (as discussed in Chapter 4) in order to be distinct from the patients used in this study. NTCP was calculated using Eq. 5.1. The resulting sigmoidal relationship between EUD and NTCP is illustrated in Figure 5.10.

$$NTCP = \frac{1}{(1+e^{-S})} \tag{5.1}$$

where S is defined as:

$$S = -19.706 + 0.302(EUD) \tag{5.2}$$

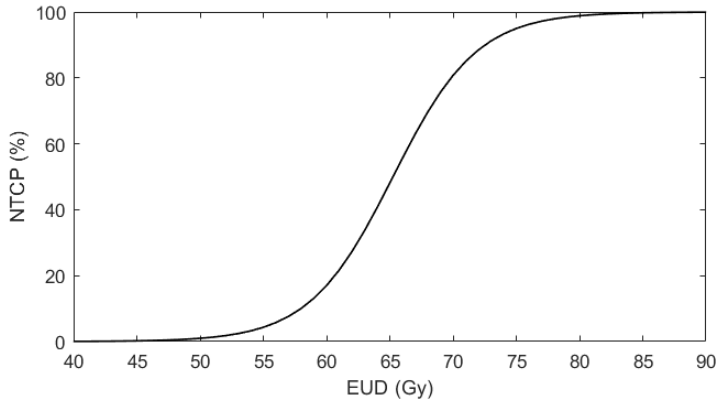


Fig. 5.10 NTCP model for rectal bleeding based on EUD, developed using univariate logistic regression on 109 patients from the discovery cohort.

5.3.3 Results

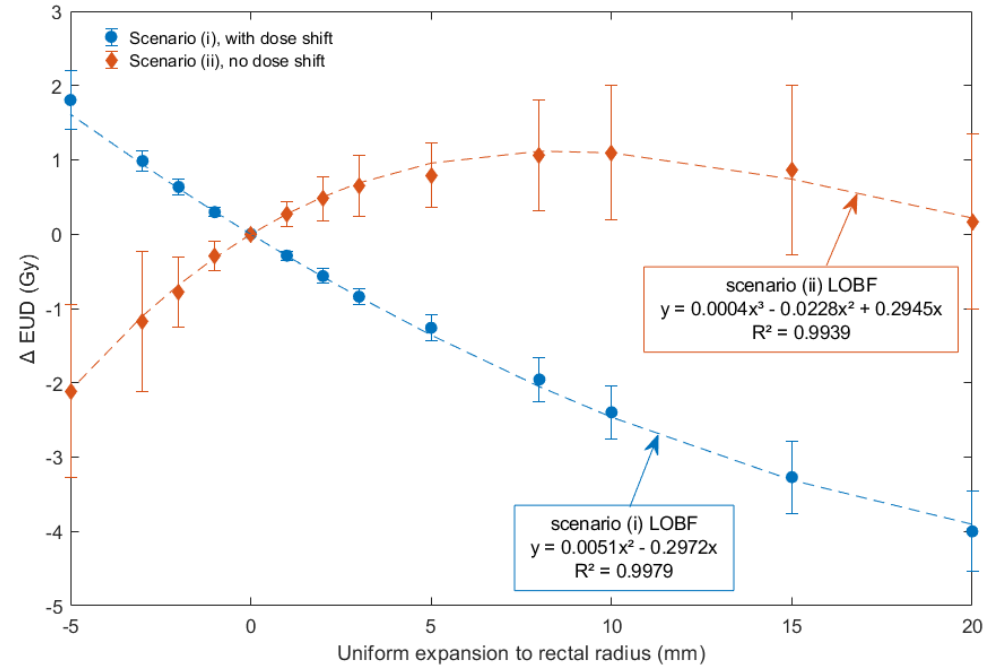
5.3.3.1 Uniform expansion

Where uniform rectal expansion induced a shift to prostate position, and hence dose positioning (scenario (i)), Δ EUD increased with rectal contraction (+1.8 Gy, SD 0.4 Gy, at -5 mm), and decreased with rectal expansion (-4.0 Gy, SD 0.5 Gy, at +20 mm), as shown in Figure 5.11a. The magnitude of change in EUD was much greater than the change observed for maximum point dose, which was less than ± 0.15 Gy from baseline across the full range of expansions from -5 to +20 mm (Figure 5.11b). Contractions less than -5 mm were not possible to simulate due to the narrow baseline contour diameters of some kVCT contours (minimum 13 mm). Across the measured range, EUD dose-response was represented using a second-order quadratic line of best fit (LOBF). Using this fit, the predicted change in EUD for a reduction in rectal diameter of -1 cm was +1.6 Gy, and an increase in diameter of +1 cm was -1.4 Gy, across the full course of treatment. This is of the order of magnitude of dose delivered in a single fraction.

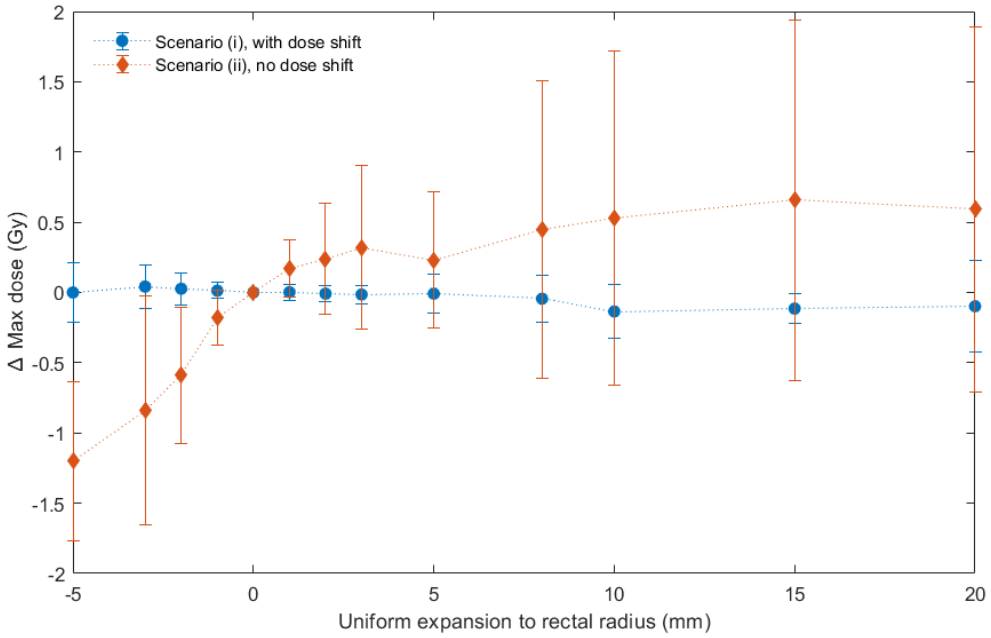
Assuming scenario (ii), where rectal expansion did not induce an anterior dose shift, the EUD dose-response is in the opposite direction to scenario (i); as shown in Figure 5.11a, an increase in diameter results in an increase in EUD with respect to baseline, and *vice versa*. The maximum increase in EUD was +1.1 Gy (SD 0.9 Gy), occurring after +2 cm increase in diameter. After this, there was a reduction in EUD which suggests that the low-dose contribution to the rectal wall begins to dominate as the maximum dose stabilises. Beyond this, the proportion of voxels receiving a lower dose increases relative to voxels receiving a high dose. The LOBF applied to the EUD dose-response was a third order polynomial. Using this fit, a reduction in rectal diameter of -1 cm corresponded to a reduction in EUD of -2.1 Gy, and +1 cm resulted in +1.0 Gy, when the change is held constant across the full course of treatment.

Maximum dose-response was of greater magnitude than scenario (i), and in the same direction as EUD response until around the point of EUD dropoff after an increase of 2 cm in diameter, where the maximum dose reaches a plateau of around 0.6 Gy (Figure 5.11b).

For each patient, the resulting EUDs from the expansion simulations for each scenario were input into the NTCP model (Equation 5.1). These were plotted in Figure 5.12, and both scenarios were modelled using third order polynomials as the LOBF. From simulated measurements, for scenario (i) based on dose shifted results, an average of 1 cm reduction in rectal diameter increased NTCP by +11.8 % (SD 3.6 %) due to a higher proportion of the rectum receiving a higher dose and hence increased NTCP. An average increase of 1 cm in rectal diameter reduced NTCP by -6.0 % (SD 2.1 %).



(a)



(b)

Fig. 5.11 Results of uniform rectal expansion simulations for (a) equivalent uniform dose (EUD) and (b) maximum point dose differences from baseline, over 37 fractions. 'With dose shift' shown in blue implies scenario (i) from Section 5.6, and 'no dose shift' shown in orange implies scenario (ii). Polynomial lines of best fit (LOBF) and corresponding R^2 are included in (a). Error bars are standard deviation of mean EUD and maximum point dose measured across all patients and increase with expansion to reflect the reduction in the number of successfully executed simulations.

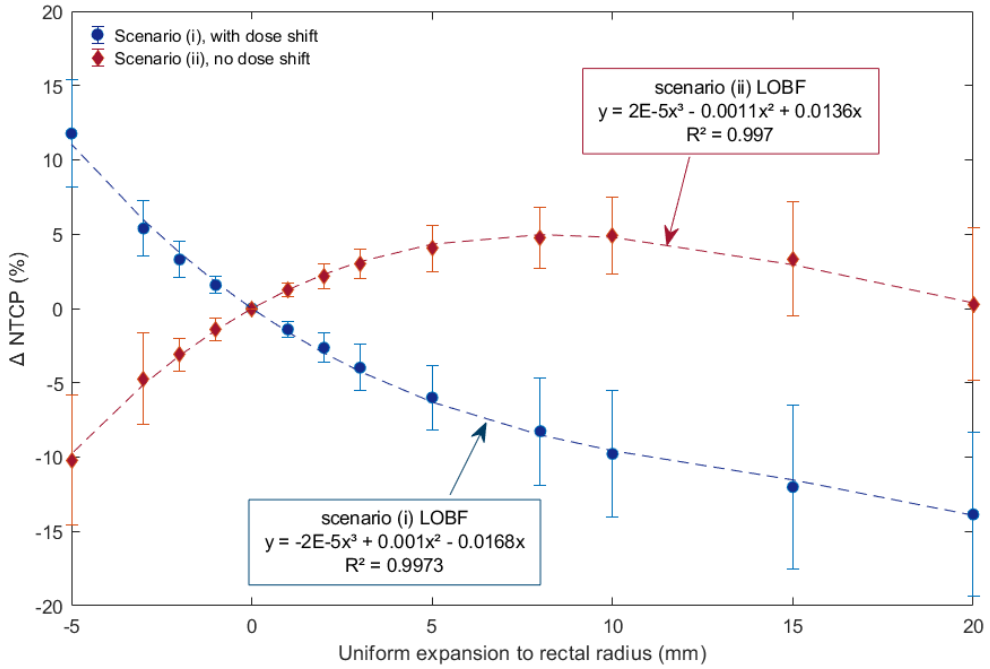


Fig. 5.12 Change in NTCP due to uniform rectal expansion scenarios based on changes in EUD over 37 fractions. 'With dose shift' shown in navy implies scenario (i) from Section 5.6, and 'no dose shift' shown in red implies scenario (ii).

Table 5.1 Predicted changes in EUD and NTCP for representative expansions to rectal diameter based on LOBF equations for uniform expansion simulations. Figures are representative assuming stated change is maintained throughout treatment.

	With dose shift, scenario (i)		No dose shift, scenario (ii)	
Δ rectal diameter	ΔEUD (Gy)	ΔNTCP (%)	ΔEUD (Gy)	ΔNTCP (%)
-1 cm	+1.8	+11.8	-2.1	-10.2
+1 cm	-1.3	-6	+0.8	+4.0
+2 cm	-2.4	-9.8	+1.1	+4.9

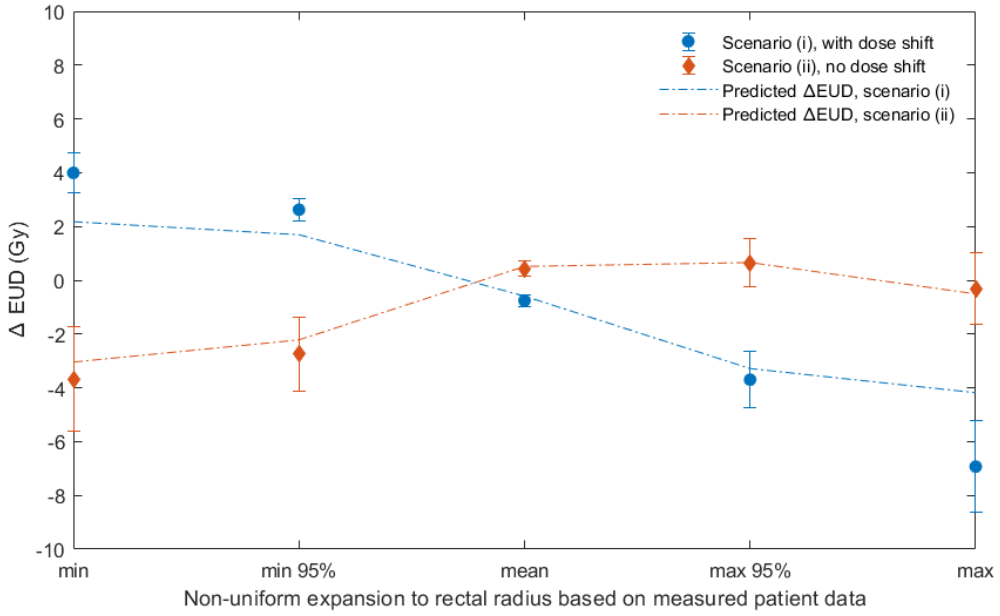
For scenario (ii) based on no dose shift, an average 1 cm reduction in rectal diameter reduced NTCP by -10.2 % (SD 4.3 %) where the rectum is contracting away from the prostate and hence high dose region. An average increase of 1 cm in rectal diameter increased NTCP by +4.0 % (SD 1.6 %), in contrast to scenario (i). The maximum increase at the point of inflection in scenario (ii) at 2 cm increase to the rectal diameter was +4.9 % (SD 1.6 %) NTCP.

5.3.3.2 Non-uniform expansion

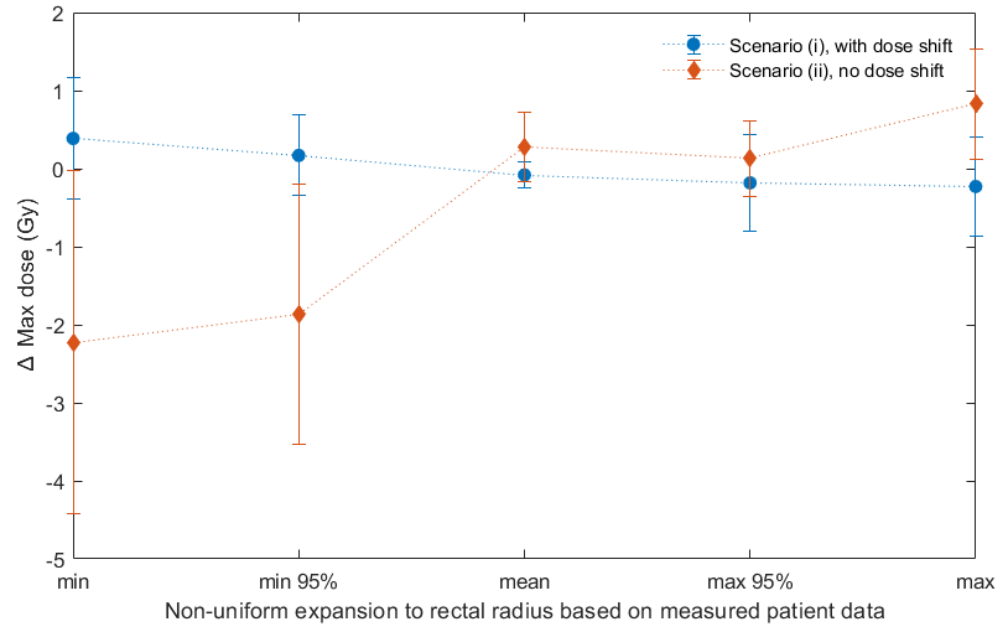
Non-uniform expansions demonstrated similar trends in dose response as results from uniform expansions (Figure 5.13). For scenario (i), where dose was shifted by the magnitude of the radial expansion at the corresponding reference slice, ΔEUD increased with rectal contraction to a maximum of +4.0 Gy (SD 0.7), and decreased with rectal expansion to a minimum of -6.9 Gy (SD 1.7). The change in maximum dose also increased with contraction, and reduced with expansion, but effects were less than $\pm$0.4 Gy. For scenario (ii) where no dose shift occurred, EUD reduced with rectal contraction to a minimum of -3.7 Gy (SD 2.0), and increased with expansion to a peak of +0.6 Gy (SD 0.9) before reducing to -0.3 Gy (SD 1.3), similar to the behaviour observed in Figure 5.11. Maximum dose changes were greater for scenario (ii), ranging from -2.2 Gy (SD 2.2) at minimum contraction, to +0.8 Gy (SD 0.7) at maximum expansion.

Predicted change in EUD based on rectal expansion was calculated using the LOBF equations derived from uniform expansion results, using the expansion at the reference slice as the input. Predicted values are represented by the dotted line in Figure 5.13a. These generally underestimate the deviation from baseline EUD (by 1.8 Gy at minimum contraction and 2.7 Gy at maximum expansion for scenario (i), and 0.6 Gy at minimum contraction for scenario (ii)). Expansion results for scenario (ii) were within 0.6 Gy of predicted values.

Changes in EUD were input into the NTCP equation 5.1 and plotted in Figure 5.14. The mean NTCP at baseline was 22.9 % (SD 10.7). The change in NTCP with dose shift, scenario (i), ranges from +23.7 % (SD 7.0) at minimum rectal contraction, to -16.1 % (SD 6.2) at maximum expansion. Where no dose shift is applied, scenario (ii), minimum rectal contraction corresponded to a reduction in NTCP of -11.2 % (SD 5.1), with a peak increase at the 95th percentile maximum expansion of +3.2 % (SD 4.8), before reducing to +0.7 % (SD 4.4) at maximum expansion.



(a)



(b)

Fig. 5.13 Results of non-uniform rectal expansion simulations for (a) equivalent uniform dose (EUD) and (b) maximum dose differences from baseline. 'With dose shift' shown in blue implies scenario (i) from 5.6, and 'no dose shift' shown in organge implies scenario (ii). Magnitudes of expansions are shown in Figure 5.8. Error bars indicate standard deviation of mean EUD and maximum point dose measured across all patients.

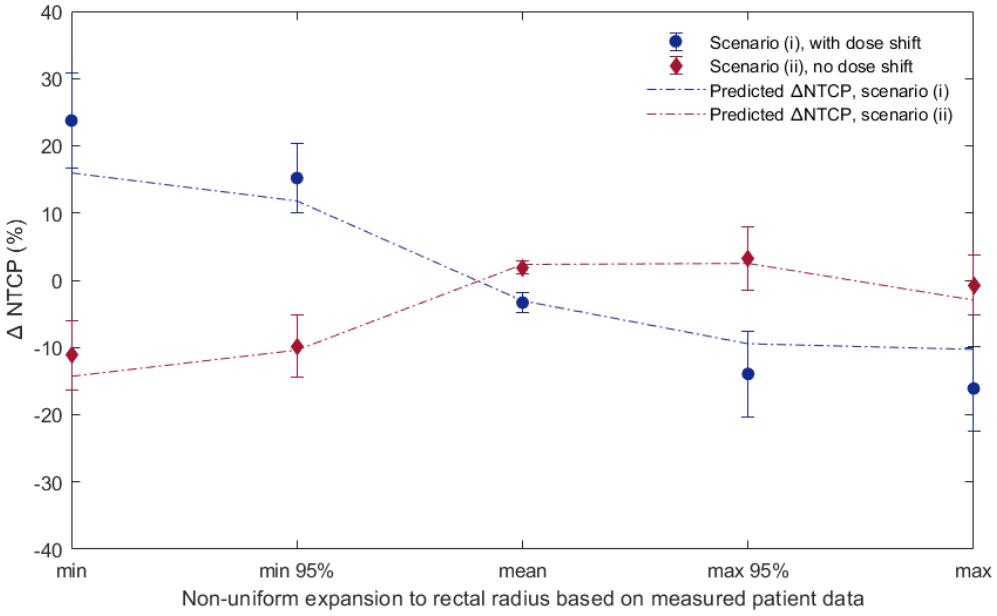


Fig. 5.14 Non-uniform rectal expansion scenarios and impact on NTCP, based on changes in EUD. 'With dose shift' shown in navy implies scenario (i) from Section 5.6, and 'no dose shift' shown in red implies scenario (ii). Error bars indicate standard deviation. Dotted lines indicate predicted change in NTCP using the line of best fit equations, where the magnitude of rectal expansion used as input into the equation was taken at the reference plane. Average NTCP at baseline was 22.9 % (SD 10.7).

Using the LOBF equations derived from uniform expansion results, the predicted change in NTCP was calculated for non-uniform expansions, based upon the change in rectal radius at the reference slice (dotted lines, Figure 5.13a). For scenario (i) where a dose shift is applied, the predicted dose underestimated the measured simulated change in NTCP (closer to baseline by 7.8 % and 5.8 % at minimum and maximum expansions, respectively).

5.3.4 Discussion

Rectal models for 26 patient were simulated in the FEA environment Abaqus based on manually segmented contours from the kVCT planning scan. Models were expanded (and contracted) either uniformly, where incremental expansions were applied to all slices of the rectal simulation, or non-uniformly, where expansions varied in the z-direction according to the measured deformations from patient data at each slice. Expansions were applied radially, but in practice the rectal diameter is measured on the IGRT image at the time of treatment, so to discuss in terms of diameter is more clinically relevant.

For each simulation, two dose scenarios were investigated; (i) where rectal expansion resulted in a dose-shift, (ii) where no dose shift was applied. These scenarios were designed to

investigate the extremes of a clinical IGRT scenario and demonstrate different dose-response behaviours. However, within each scenario the direction of the dose and NTCP responses agreed for uniform and non-uniform expansions.

Scenario (i) suggests that where rectal expansion causes the prostate to move, and the dose to be shifted accordingly, that rectal expansion causes EUD to reduce, and rectal contraction causes EUD to increase. Hence, a narrower rectal diameter with respect to the planning scan would be associated with an increased likelihood of developing toxicity. This result implies that achieving as narrow a rectal diameter as possible at the time of the kVCT planning scan is therefore desirable, as further narrowing throughout the course of treatment would be less probable. On-treatment clinical imaging protocols often define tolerances for the maximum acceptable rectal diameter, but do not specify a minimum diameter. Furthermore, often patients will be instructed to empty their bowels if the rectum is deemed outwith the maximum tolerance. Results for scenario (i) suggest this practice may need revising, as larger rectal diameters are associated with decreased dose and NTCP, whereas narrower rectums should be monitored. Thor et al. [155] reported findings that narrower rectums on the planning CT scan were associated with higher rates of morbidity. However, VoxTox results presented in Section 5.2 showed that kVCT diameter was not indictive of MVCT diameter, and hence delivered dose. Maximum dose appeared to be fairly independent of rectal wall expansion or contraction, varying by less than ± 0.4 Gy.

The results for scenario (ii) are more intuitive, whereby the EUD increases as the rectum expands into the static high-dose region, and reduces as the rectum contracts away from it. The EUD reaches a peak and begins to reduce as the maximum dose begins to plateau. This response suggests that at this point, as the rectal radius increases, the proportion of the rectum receiving a lower dose is increasing greater than the proportion of the rectum receiving a higher dose, resulting in a reduction in EUD.

The direction of the dose response for each scenario agreed between uniform and non-uniform expansions. For scenario (i), rectal expansion resulted in a reduction in dose and NTCP. This suggests that with an increase in rectal diameter, the proportion of rectal wall receiving a high dose level is reduced, and the risk of toxicity is lowered. Rectal contraction resulted in an increase in dose and NTCP, at a greater rate than expansion results. For uniform expansion with a corresponding dose shift, dose response was modelled using polynomials and can be interpreted for clinical purposes as approximately: a reduction of -1 cm in rectal diameter results in an increase of +1.5 Gy (or +10 % NTCP), and an increase in rectal diameter of 1 cm corresponds to a reduction of -1.5 Gy (or -5 %), where the rectum remains abutting the prostate maintained over the course of treatment. The results

for non-uniform expansion suggest that these may be under-estimations of the magnitude of the dose/NTCP-effects.

When comparing simulated dose measurements for non-uniform rectal expansions with predicted dose measurements from the equations derived for uniform rectal expansion simulations for scenario (i), the predicted values underestimate the absolute effect on EUD, but are within the bounds of the standard deviation error bars (with the exception of minimum expansion). The differences may be due to expansions only being applied to the central 15 slices, and slices outwith this range, where the dose is equal to the original planning dose, may be negatively influencing the EUD calculation. Additionally, radial expansions used in the prediction calculation, were from the reference slice, but the magnitude of the expansion at each slice varied, which was not accounted for in the prediction model.

For scenario (ii), the trend of the dose response with expansion and contraction from uniform simulations were generally in agreement with non-uniform simulations. The changes in EUD were in the opposite direction when compared to scenario (i), with EUD increasing with rectal expansion, and reducing with rectal contraction. EUD reached a maximum when the rectal diameter had increased by approximately 2 cm. However, this situation is fairly clinically unrealistic. The magnitudes of the maximum absolute deviations from baseline EUD were less than the dose effects in scenario (i).

Maximum dose demonstrated more variability in scenario (ii) than scenario (i), deviating from baseline by -2.2 to + 0.8 Gy for non-uniform simulations. The upper limit of the maximum dose is determined by the planned dose prescription, and so such an increase is greater than expected given the fairly homogeneous doses delivered to the PTV. For minimum contraction, where the rectum moves away from the static high-dose region, the fall-off in maximum dose observed is as expected.

Predicted EUD and NTCP response based on uniform expansion underestimated the simulated changes measured for non-uniform expansions in the case of a dose shift being applied. In this scenario (i), a wider rectal diameter results in lower EUD and NTCP, and a narrower diameter results in a higher EUD and NTCP. The expansion parameters used in the equation were based upon expansions at the reference slice which were greater than the expansions recorded at most other slices (i.e. rectum narrowed with increasing distance from reference slice), so should lead to an overestimation in ΔEUD and ΔNTCP. Although the predicted values generally lie within the standard deviation error boundaries, it is unclear why these are different than expected. Perhaps the use of a single uniform expansion coefficient along the length of the rectum may not be adequately physiologically representative, and could limit the accuracy of the predictive model. As shown in Figure 5.5, the superior rectum is more susceptible to motion than the inferior rectum so perhaps further investigation into

typical range and behaviour of motion within superior, middle, and inferior rectal thirds may improve prediction of the non-uniform range of motion measured throughout the course of treatment. For scenario (ii), where no dose shift is applied, the predicted change in EUD for rectal contraction is underestimated with respect to baseline, yet the NTCP effect is overestimated. However, expansion predictions lie very close to measured for EUD and NTCP, and all predicted results lie within one standard deviation of the measured result.

The dosimetric implications from this FE study can also be interpreted as a simulation of contouring sensitivity; for instance, an inter-clinician discrepancy of 2 mm in defining the radius of the rectal wall might induce a dose difference of ± 0.6 Gy for the same patient anatomy, which corresponds to a change in NTCP of approximately $\pm 3\%$.

Several limitations to this study should be identified. Firstly, not all simulations could be completed by the FEM system. For the more extreme scenarios being simulated for uniform expansion, an increasing number of simulations failed to produce analytical solutions as the magnitude of simulated radial expansion increased. This is represented within the standard deviations quoted alongside the results, which increase with absolute expansions. Secondly, in reality, rectal expansion and contraction are non-uniform along the length of the rectum in the patient z-axis, so predictions based on uniform expansion may not be representative (and were shown to underestimate the dose-response of the clinical situation). However, it is proposed that predictions may be adequate, with acknowledged caveats, for use as a 'rule-of-thumb' guide. Validation of the prediction equations on a separate patient cohort would be beneficial. To independently test model generation would require planned EUD to the rectal wall combined with late toxicity data, or to assess prediction equations would require planning structures and planned dose as input into FE simulations. This could potentially be achievable using data in the literature [3, 35, 52]. An additional limitation of the study relates to the use of EUD, a single metric describing a 3D dose distribution. Although selected for its simplicity, results demonstrate a non-linear relationship between EUD-response and maximum dose response, assumed to be caused by the relative increase in low dose contribution after a certain expansion threshold (which appears to occur at approximately +2 cm in diameter). This effect may be influencing the NTCP calculation which does not consider absolute maximum dose. Finally, the stated changes in EUD and NTCP are reported in terms of entire treatment prescription and are based on the assumption that a given expansion is maintained throughout the course of treatment. This is unlikely in practice, but nevertheless, it is important to appreciate how rectal deformation impacts dose, and consequently NTCP, based on the full range of measured rectal motion.

To summarise, the effect of rectal motion on dose has been investigated using biomechanical FE modelling. Gradually increasing intervals of uniform expansion and contraction were

applied to investigate sensitivity, and non-uniform expansions and contractions were applied throughout the length of the rectum to model the extremes of measured clinical contours. The behaviour of the dose response depends on whether corrective dose shifts are applied or not. This was represented by two different clinical scenarios when registering dose. Assuming scenario (i) where dose is shifted by the magnitude of the rectal expansion, a narrower rectal diameter at the time of treatment indicates that the delivered dose to the rectum will be greater. Hence, a constricted rectum at treatment indicates a greater probability of rectal bleeding. These results emphasise the importance of coherence with rectal filling protocols at treatment planning, whereby, minimising the rectal diameter at treatment planning would be representative of a worst-case dosimetric scenario, and therefore delivered dose to rectum would be less than planned if the rectal diameter at treatment was greater than at planning. Rectal filling protocols in IGRT generally advise a maximum dimensions that the rectal diameter should not exceed. However, these results suggest that there should be a tolerance on the minimum rectal diameter at treatment with respect to baseline. In scenario (ii) where rectal expansion results in a greater proportion of the rectal wall encroaching into the high dose region, reactive measures should be taken (e.g. repositioning the patient). Future work could include contouring of the prostate on MVCT image guidance scans to determine whether scenario (i) or (ii) is more common.

In terms of recommendations for existing clinical practice, the results suggest that if the rectal diameter at the time of treatment is consistently smaller than the diameter measured at treatment planning, in the region where the rectum is abutting the PTV, a dosimetric evaluation should be undertaken, with a view to adapting the plan to minimise the potential risk of rectal toxicity.

Chapter 6

Biomechanical modelling for dose-toxicity prediction

In this chapter, biomechanical finite element (FE) modelling is introduced to the VoxTox process with the aim of improving accuracy and robustness of dose calculation. The simulated rectal model facilitates tracking of dose-histories at the voxel level throughout the course of treatment. Voxels of the rectal wall deform in 3D according to biomechanical properties and applied boundary conditions, providing a more anatomically representative basis for dose accumulation, intra- and inter-patient. The approach described in Chapter 4 for two-dimensional dose-surface maps (2D-DSMs) was limited to in-plane, isotropic expansion only, where voxels were uniformly normalised around the circumference. Calculating and accumulating dose to the rectal wall using FE generated dose-surface maps (FE-DSMs) is a novel approach that increases the resolution of dose information available for interrogation. Here, we investigate different methods for parameterising dose from FE-DSMs and explore dose-toxicity associations.

I designed and conducted the work described in this chapter. My role was to integrate existing FE code into the VoxTox workflow, edit the code so that it was able to handle patient DICOM data, and optimise the model for patient conditions rather than experimental conditions. I wrote the code for generating geometric subregions, and devised the approach for determining statistical subregions at risk using binary masks. I processed toxicity data and combined with baseline and dosimetric information, including careful division into training and validation sets for each endpoint. All dose-toxicity analysis (univariate and receiver operator characteristic analysis) was performed by me.

I would like to acknowledge the contributions of Dr Simon Thomas who wrote the code to generate 2D-DSMs and provided guidance for combining 60 Gy and 74 Gy fractionation regimes, Dr Marina Romanchikova who provided assistance in generating the code for

dose-widths and EUD calculation, Prof Michael Sutcliffe who wrote the code for performing FE analysis, and Dr David Noble, Mrs Amy Bates and Mr Karl Harrison who played vital roles in the extraction and interpretation of toxicity data.

6.1 Introduction

A biomechanical model is a physical-based model that incorporates the mechanical properties of soft tissue in addition to geometry and boundary conditions [21]. Finite element (FE) modelling has been adopted for anatomical modelling within the radiotherapy community for purposes such as deformable image registration for intra- and inter-imaging modality, improving target localisation, and even dose accumulation [20, 21, 114, 135, 163, 165]. The novelty and advantage offered by FE analysis is that the material elements, or voxels, are able to deform in a more anatomically realistic way [21], according to constitutive biomechanical laws. FE analysis provides a sophisticated tool for tracking organ motion, which facilitates the ultimate aim of the VoxTox project - to track dose at the voxel level.

A limitation of the 2D-DSM method presented in Chapter 4 was that rectal expansion was assumed to be restricted to the plane of the transverse imaging slice, as well as being isotropically normalised around the circumference. The advantage of voxel-tracking using FE analysis is that anisotropic, extraplanar expansion is possible. FE analysis is even beginning to be recognised by manufacturers of treatment planning systems as a important feature in dose accumulation [165]. By implementing FE modelling into the generation of DSMs within the VoxTox process, we can achieve a more anatomically representative registration and improve the accuracy in dose calculation and accumulation.

An alternative voxel-based solution for registration of the rectum has been previously reported in the literature, whereby the planned dose to the rectal wall for 118 patients were combined to a common anatomical template [59, 61]. By splitting the resulting 3D voxel-based dose-maps into sub-regions at risk (SRRs) based on geometric and probabilistic segmentation, features of the planned DVH of each SRR were investigated for associations with toxicity. This SRR approach was adapted for our work, whereby: geometric subregions were divided into anterior/posterior hemispheres, anterior/posterior/left/right longitudinal quarters, and superior/middle/inferior lateral bands around the circumference; and statistical subregions were determined by performing independent sample t-tests on patients with and without toxicity.

Here, we test the hypothesis that extracting dosimetric parameters from FE-DSMs improves the strength of dose-toxicity association. As well as comparing FE and 2D-DSM

analysis methods, we compare different parameterisation techniques of increasing complexity. The chapter is divided into three sub-studies:

- FE-DSM dose-widths and equivalent uniform dose (EUD), where parameters extracted from FE-DSMs are compared with previous results for 2D-DSMs
- Spatial division of FE-DSMs into geometric subregions. EUD is calculated for each subregion and compared with EUD of the full rectal wall.
- Statistically derived subregions at risk are determined by examining dose distributions of patients with and without toxicity.

Dose-toxicity analysis is performed for each analysis and parameterisation approach. Results are then compared between sub-studies to interpret the benefit of dose accumulation using FE-DSMs for toxicity prediction.

6.2 Parameterising FE-DSMs using dose-widths and EUD

6.2.1 Materials and methods

Using the system described in Chapter 3.6, the data from 139 VoxTox prostate cancer patients from the training set of the consolidation cohort were processed using FE-DSMs to present planned dose and accumulate delivered dose. Patients were included where a minimum of 2 years' follow-up data were available, acquired at 6, 12, and 24 months ($\pm$2 months) post-treatment. Baseline data and details of clinical cofactors were available for all patients in the consolidation cohort. Patients were prescribed either 74 Gy in 37 fractions (59%), or 60 Gy in 20 fractions (41%), and were treated with helical image guided radiotherapy (IGRT) on TomoTherapy. Twelve toxicity endpoints were investigated (including different staging levels), as detailed in Table 6.1.

For one of the patients, the FEA model could not be successfully completed using the planning data. This was due to excess distortion at one of the nodes. In practice, this could be due to a large rectal diameter at the time of CT scanning, and could be solved in the future by stricter adherence to restrictions imposed within imaging protocols, or alternatively, may be due to an erroneous contour generated by the autocontouring system. Consequently, baseline planned dose information was missing so this patient was removed from the analysis. Toxicity breakdowns are shown for the resulting 139 patients in each dataset in Table 6.1. Excess nodal distortion also caused the model to terminate prematurely for 1-2 fractions for 4 patients, resulting in missing daily dose data. Due to the relatively low number of fractions

Table 6.1 Consolidation cohort toxicity incidence for 139 patients in the training set.

Toxicity	Scoring System	Incidence % (n)
Diarrhoea $\geq$ G1	CTCAE [158]	25.9 (36)
Faecal incontinence $\geq$ G1	CTCAE [158]	17.3 (24)
Proctitis $\geq$ G1	CTCAE [158]	18.7 (26)
Proctitis $\geq$ G2	CTCAE [158]	14.4 (20)
Rectal bleeding $\geq$ G1	CTCAE [158]	33.1 (46)
Rectal bleeding $\geq$ G2	CTCAE [158]	11.5 (16)
GI toxicity $\geq$ G1	RTOG [127]	52.5 (73)
GI toxicity $\geq$ G2	RTOG [127]	15.8 (22)
Stool frequency $\geq$ G1	LENT SOMA [1]	34 (24.5)
Bowel bother $\geq$ G1	UCLA-PCI [89]	38.8 (54)
Bowel bother $\geq$ G2	UCLA-PCI [89]	22.3 (31)
Bowel bother $\geq$ G3	UCLA-PCI [89]	9.4 (13)

affected, missing daily delivered dose data were substituted with planned dose data for the number of affected fractions.

For each patient, rectal models were generated for planned data, as well as for each of the 20 or 37 daily treatment fractions. Image sets and contours were processed in Matlab to determine coordinate systems, positional offsets, contour locations, and centrelines. The data were consolidated into input files for the Abaqus simulation as described in Sections 3.6 and 5.1.

A central mandrel, based on the planning contour, was used as the common starting geometry for each daily fraction rectal model so that the material elements, or voxels, can be tracked. The model expands and deforms to the daily MVCT contour according to the biomechanical properties and boundary conditions applied. The model used was a hyperelastic neo-Hookean, appropriate for rubber-like soft tissues. Model properties were previously described in Chapter 3.6.

A Python script was run to extract surface coordinates of the deformed model from Abaqus. Delivered dose to each voxel was then calculated via interpolation of the daily delivered dose matrix in MatLab. Where the patient's fractionation schedule was 60 Gy in 20 fractions, planned and delivered doses were converted to the equivalent dose in 37 fractions, using an α/β ratio of 2.149 (as detailed in Section 3.3).

To compare the FE-DSM approach with the 2D-DSM approach discussed in Chapter 4, the same dose parameters were investigated. DSM 'dose-widths' - the lateral extent of an ellipse fitted to a discrete isodose level - were measured for doses 30 - 70 Gy in 5 Gy increments, and EUD - a form of weighted mean dose across the full DSM - were calculated for FE-DSMs.

These parameters were extracted and compared for planned and accumulated FE-DSMs, and were independent of toxicity endpoint. For each of the 12 endpoints investigated, the composition of individual patient datasets differed so that toxicity incidence rates could be retained. However, this difference is not expected to significantly affect results comparing FE-DSM vs. 2D-DSM or planned vs. accumulated dose. For our VoxTox patient population, the patients used as the representative sample for this analysis was from the rectal bleeding dataset (n = 139).

For each of the 12 endpoints, a comparison was made between the mean FE-DSMs of planned and accumulated dose, patients with and without toxicity for both planned and accumulated dose, and the difference between planned and accumulated toxicity difference maps. The greatest dose differences observed in the former indicates the areas most susceptible to interfraction motion, and in the latter indicates the geometric areas where additional information on toxicity prediction might be available from accumulated dose.

For a comparison with the 2D-DSM results (Chapter 4), receiver operator characteristic (ROC) curve analysis was performed on the unfolded FE-DSM for EUD, and dose-widths of 30 - 70 Gy in 5 Gy intervals. To further investigate spatial aspects of dose, the unfolded FE-DSMs were divided geometrically into rectal subregions at risk (SRR) following the methods of Dréan et al. [61]. SRRs were: anterior (ant) and posterior (post) hemis; ant, post, right, left quarters, split longitudinally; superior (sup), middle (mid), inferior (inf) thirds, split laterally; and a combination of lateral splits divided by longitudinal splits. Finally, SRRs were selected based on statistical analysis of the difference maps per endpoint.

Planned and accumulated delivered dose were calculated using FE analysis for a dataset of 139 patients, distinct for each of the 12 endpoints. FE-DSMs were generated by cutting the 3D model posteriorly, and virtually unwrapping to a 30×80 element 2D map. The cutting point was defined as the most posterior point of the starting central mandrel, which was common to each slice.

6.2.2 Results

6.2.2.1 FE-DSM vs. 2D-DSM

FE-DSM parameters were compared with parameters extracted from 2D-DSMs for the same patients, which were previously presented in Chapter 4. Dose-widths measured from FE-DSMs were greater than those measured from 2D-DSMs across all dose levels investigated, ranging from an average increase of 2.1 % (SE 0.2 %)/2.3 % (SE 0.2 %) for planned/accumulated 30 Gy dose-widths, to 7.0% (SE 0.4 %)/8.6 % (SE 0.4 %) for planned/accumulated 70 Gy dose-widths. Figure 6.1 shows how the average difference

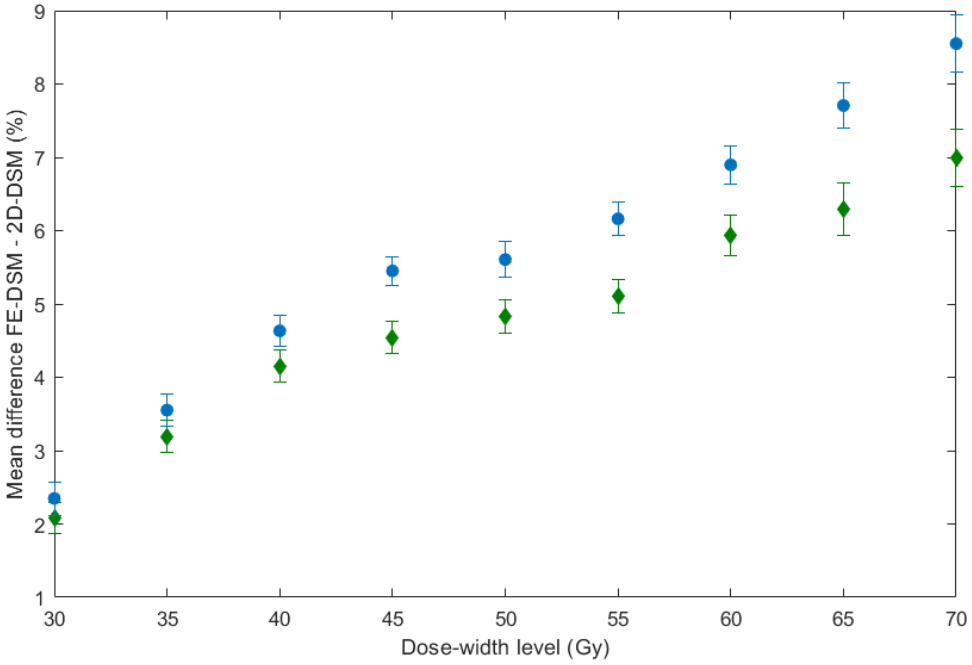


Fig. 6.1 Mean difference between FE-DSM and 2D-DSM dose-widths for planned (green) and accumulated (blue) dose. Error bars indicate standard error measured across 139 patients in the rectal bleeding training cohort. Accumulated dose differences are higher, and increase with dose-width.

between FE-DSM and 2D-DSM calculated dose-widths increases with dose level. EUD was also greater when calculated using FE-DSMs with respect to 2D-DSMs for both accumulated dose, with a mean increase of 0.68 Gy (SE 0.05 Gy), and planned dose, with a mean increase of 0.19 Gy (SE 0.03 Gy). The differences between FE-DSM and 2D-DSM calculated dose were significant (paired samples t-test, $p<0.001$) for all measured dose parameters.

6.2.2.2 Planned vs. accumulated dose

The difference between mean planned and mean accumulated FE-DSMs is shown in Figure 6.2. The difference map shows an inhomogeneous distribution, emphasising the need for spatial dose quantifiers. Analysing FE-DSMs, the average accumulated EUD was -1.1 Gy (SE 0.1 Gy, range [-6.6, 1.3] Gy) with respect to planned EUD. Differences between planned and accumulated EUD were significant ($p < 0.001$). Accumulated EUD was lower than planned for 112 of 139 patients, with a difference of -1 Gy or lower observed for 69 patients and -2 Gy or lower for 22 patients. Differences between planned and accumulated dose-widths were significant at intermediate to higher dose levels from 40 - 70 Gy ($p<0.01$), but not at 30 and 35 Gy dose-widths. On average, accumulated dose-widths were lower than planned dose

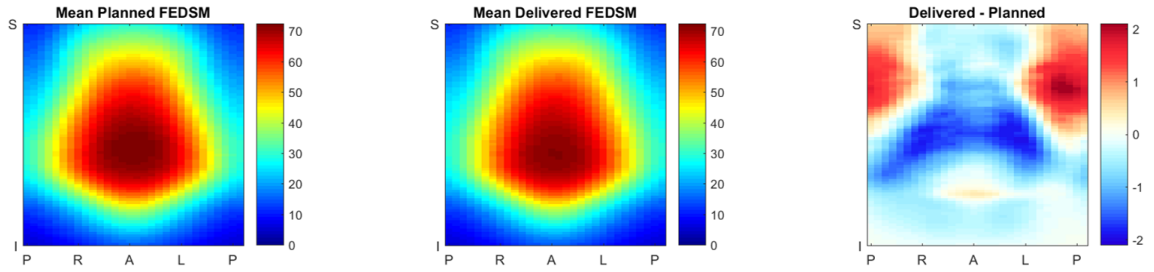


Fig. 6.2 Spatial dose differences between mean planned and mean accumulated FE-DSMs for 139 patients. Note inhomogeneous distribution of difference map.

widths for all dose levels above 30 Gy, and the magnitude of the difference increased with dose level.

6.2.2.3 Dose-toxicity analysis

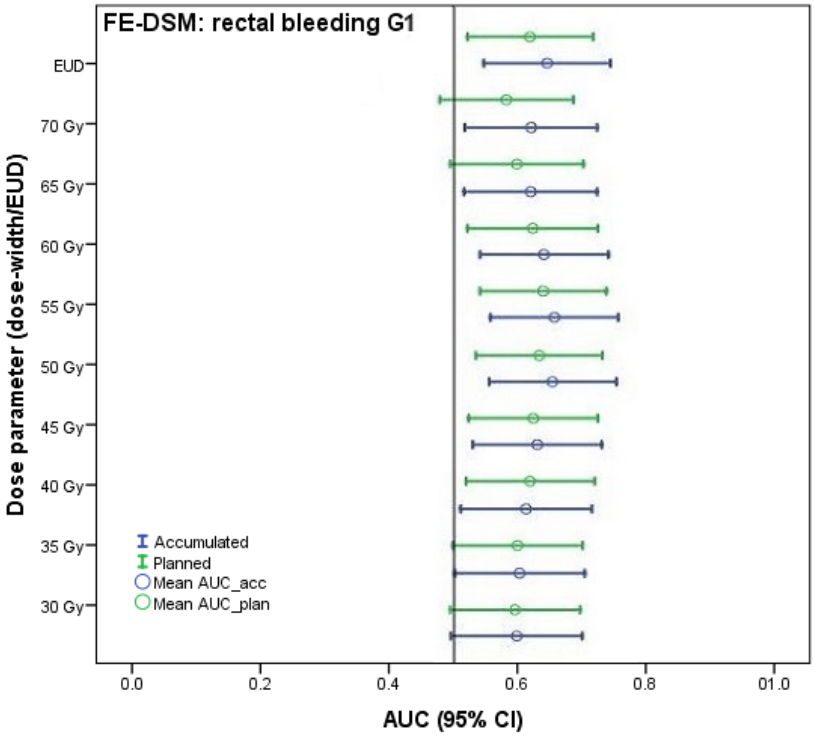
Results of dose-toxicity analysis for all endpoints are discussed below and the most discriminative dose metrics are listed in Table 6.2. The area under the receiver operator characteristic curve (AUC) and 95 % confidence intervals (95%CI) for rectal bleeding $\geq$ G1 and $\geq$ G2, RTOG GI toxicity $\geq$ G2, and stool frequency $\geq$ G21are shown in Figure 6.3.

Diarrhoea (n = 36, 25.9%)

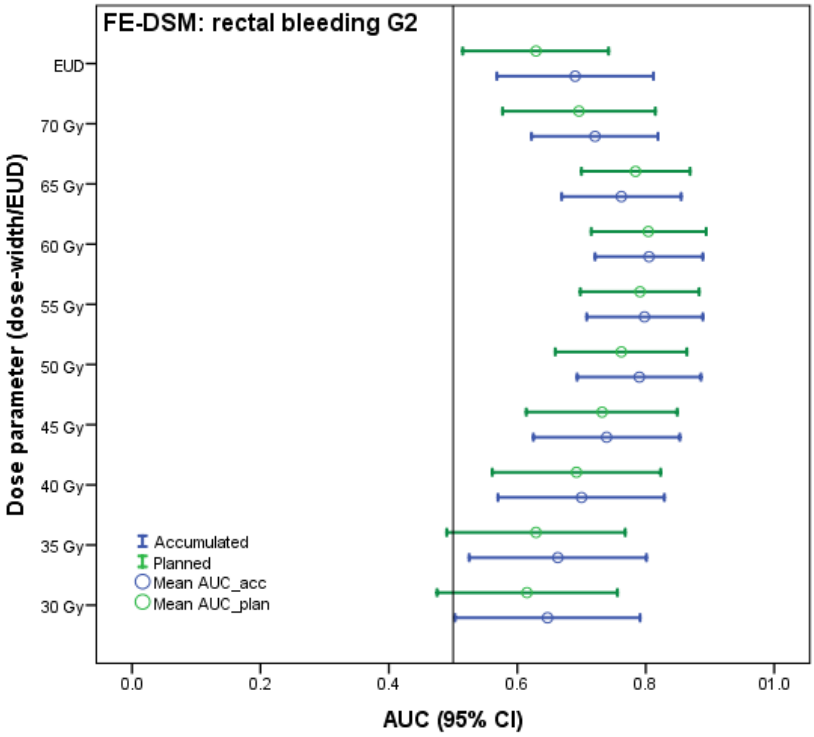
Planned dose-widths from 30 to 45 Gy were predictive of diarrhoea, with the planned 35 Gy dose width producing the largest AUC of 0.654 [95% CI 0.550, 0.759]. The mean difference in planned 35 Gy dose-width between patients with and without diarrhoea was 4.1 % (p = 0.022). Accumulated dose-widths from 35 to 55 Gy and EUD were also predictive. Similar trends were observed when compared with 2D-DSM results, where planned 35 Gy dose-width was the strongest predictor (AUC 0.689, [95% CI 0.587, 0.785]), and associations were observed for planned dose-widths from 30 - 50 Gy, accumulated dose-widths from 30 - 60 Gy, and accumulated EUD.

Faecal incontinence (n = 24, 17.3%)

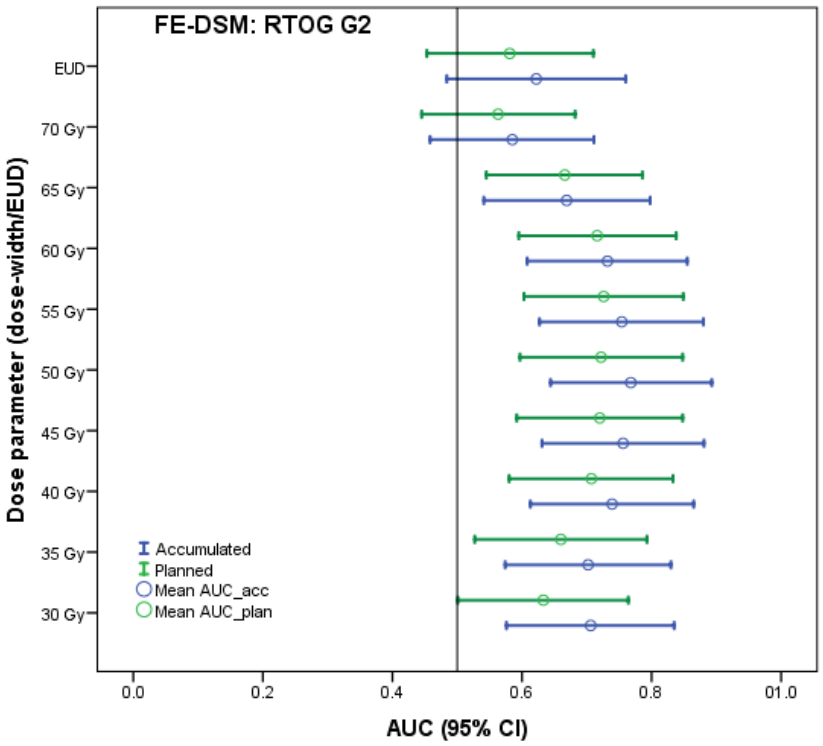
Planned and accumulated dose-widths from 40 to 55 Gy, and planned 60 Gy dose-widths, produced AUCs greater than 0.6. However, the lower bounds of the 95%CI dropped below 0.5 so no single dosimetric parameter was considered a significant predictor of faecal incontinence. The same result was found from 2D-DSM analysis.


FE-DSM: rectal bleeding G1
Dose parameter (dose-width/EUD)
EUD
70 Gy
65 Gy
60 Gy
55 Gy
50 Gy
45 Gy
40 Gy
35 Gy
30 Gy
Accumulated
Planned
Mean AUC_acc
Mean AUC_plan
0.0
0.2
0.4
0.6
0.8
01.0
AUC (95% CI)

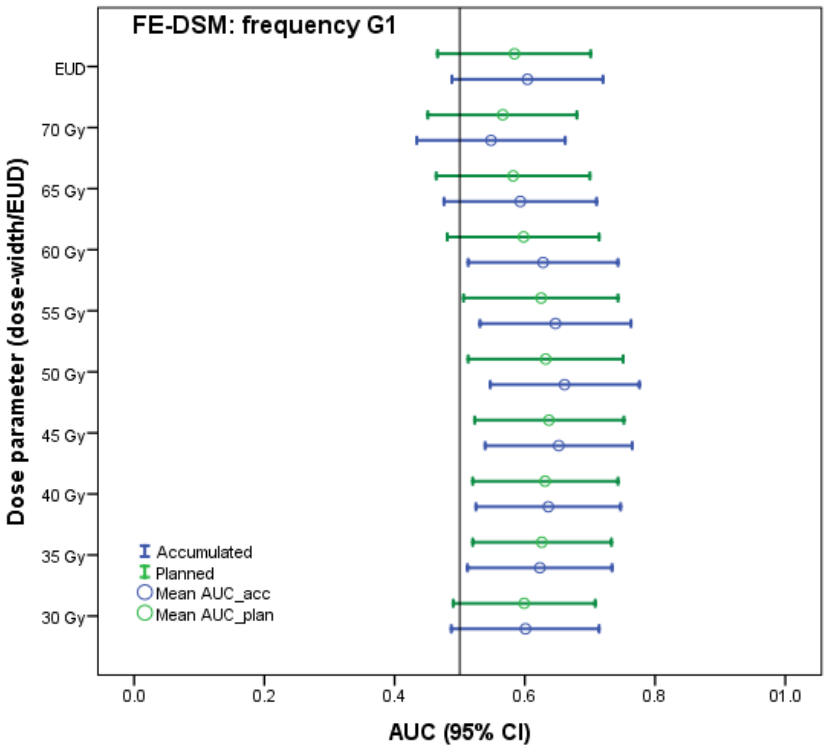
(a)


FE-DSM: rectal bleeding G2
Dose parameter (dose-width/EUD)
EUD
70 Gy
65 Gy
60 Gy
55 Gy
50 Gy
45 Gy
40 Gy
35 Gy
30 Gy
Accumulated
Planned
Mean AUC_acc
Mean AUC_plan
0.0
0.2
0.4
0.6
0.8
01.0
AUC (95% CI)

(b)



(c)



(d)

Fig. 6.3 Area under the receiver operator characteristic curve (AUC) and 95% confidence intervals (CI) for planned (plan) and accumulated (acc) dose parameters (dose-widths and equivalent uniform dose, EUD), for four of 12 toxicity endpoints investigated: (a) Rectal bleeding $\geq$ G1, (b) Rectal bleeding $\geq$ G2, (c) RTOG GI toxicity $\geq$ G2, (d) Stool frequency $\geq$ G1. Endpoints were selected as examples where FE-DSM analysis improved upon dose-toxicity associations found from 2D-DSM analysis, and where parameters from accumulated dose were overall greater predictors than from planned dose.

Proctitis $\geq$ G1 (n = 26, 18.7%)

For proctitis $\geq$ G1, the strongest predictor was the planned 45 Gy dose-width (AUC 0.688, [95% CI 0.567, 0.809]). The mean difference between patients with and without toxicity was 6.0 % (p = 0.002). Planned EUD and dose-widths from 30 to 65 Gy were also discriminative. Accumulated dose widths at 60 and 65 Gy were predictive of proctitis $\geq$ G1, but produced AUCs lower than corresponding planned dose-widths. From 2D-DSM analysis, the strongest predictor was the planned 65 Gy dose-width with AUC = 0.650 [95% CI 0.547, 0.753].

Proctitis $\geq$ G2 (n = 20, 14.4%)

Planned 40 and 45 Gy dose-widths were strongest predictors of proctitis $\geq$ G2, both with AUC = 0.674, and 95%CIs [0.540, 0.808], [0.539, 0.808], respectively. Differences between planned dose-widths for patient with and without G2 proctitis were 6.3 % (p = 0.007) and 5.8 % (p = 0.009) for 40 Gy and 45 Gy, respectively. Planned dose-widths from 30-60 Gy were also predictive. Accumulated dose-widths at 30, 40-65 Gy were associated with G2 proctitis. AUCs were greater from planned dose-widths up to 55 Gy, and accumulated AUCs were greater for 60 and 65 Gy. This result is in contrast to the 2D-DSM result which found the strongest predictor to be the accumulated 60 Gy dose-width with AUC = 0.648 [95% CI 0.524, 0.772].

Rectal bleeding $\geq$ G1 (n = 46, 33.1%)

The strongest predictor of $\geq$ G1 rectal bleeding was the accumulated 55 Gy dose-width, AUC = 0.652 [0.555, 0.749], as shown in Figure 6.3a. The difference in accumulated 55 Gy dose-width between patients with and without toxicity was 4.6 % (p = 0.003). Accumulated EUD and dose-widths from 40-70 Gy were also associated with $\geq$ G1 rectal bleeding, as well as planned EUD and 40-60 Gy dose-widths. At the 30 Gy dose-width, planned AUC was greater than accumulated, and for EUD and dose-widths above 45 Gy, accumulated AUCs were greater than planned AUCs. The 2D-DSM analysis results also found accumulated dose to be generally more predictive than planned dose, but the strongest predictor was accumulated EUD with AUC = 0.651 [95% CI 0.556, 0.747].

Rectal bleeding $\geq$ G2 (n = 16, 11.5%)

Accumulated EUD and 30-70 Gy dose-widths were associated with $\geq$ G2 rectal bleeding, as shown in Figure 6.3b, with the strongest predictor being the 60 Gy dose-width (AUC = 0.805 [0.721,0.889]). The difference in accumulated 60 Gy dose-width between patients with and

Table 6.2 Dosimetric parameters most discriminative of toxicity from planned and accumulated dose surface maps generated using finite element modelling. Dose-width levels or equivalent uniform dose (EUD) resulting in the greatest area under the receiver operator characteristic curve (AUC) are listed. Endpoint entries were omitted where no parameter from planned or accumulated dose produced an AUC with minimum 95% confidence interval greater than 0.5, so were not considered significant. A dash entry in the table indicates that no significant parameter was determined. Bold indicates greatest AUC between planned and accumulated results.

	Planned dose		Accumulated dose	
Toxicity endpoint	Dose-width (Gy)/EUD	AUC	Dose-width (Gy)/EUD	AUC
Diarrhoea	35	**0.654**	50	0.623
Proctitis ≥G1	45	**0.688**	65	0.642
Proctitis ≥G2	40	**0.674**	60	0.652
Rectal bleeding ≥G1	55	0.635	55	**0.652**
Rectal bleeding ≥G2	60	0.804	60	**0.805**
GI toxicity ≥G2	55	0.726	50	**0.768**
Stool frequency	45	0.637	50	**0.661**
Bowel bother ≥G1	45	**0.616**	60	0.611
Bowel bother ≥G2	-	-	EUD	**0.615**

without G2 rectal bleeding was 8.4 % ($p < 0.001$). Planned dose-widths at 40-70 Gy and EUD were also correlated with ≥ G2 rectal bleeding. AUCs were greater from accumulated dose than planned dose for EUD, and all dose-widths investigated, except the 65 Gy dose-width. From 2D-DSM analysis, the planned 65 Gy dose-width was most discriminative, with AUC = 0.793, [95% CI 0.700, 0.885].

RTOG GI toxicity ≥ G1 (n = 73, 52.5%)

There were no dosimetric parameters from planned or accumulated FE-DSMs that produced AUCs greater than 0.6.

RTOG GI toxicity ≥ G2 (n = 22, 15.8%)

The strongest predictor of RTOG GI toxicity ≥ G2 was the accumulated 50 Gy dose-width, with AUC 0.763 [95% CI 0.644, 0.893], as shown in Figure 6.3c. The difference between patients with and without toxicity was 7.4 % ($p < 0.001$). At dose-widths from 30-65 Gy, both accumulated and planned dose were significant predictors. Accumulated dose produced greater AUCs than planned dose for all measured dose parameters. Results mirror those from the 2D-DSM analysis, which also found the accumulated 50 Gy dose-width to be most predictive, with AUC = 0.743 [95% CI 0.621, 0.865].

Stool frequency (n = 34, 24.5%)

Accumulated dose-widths from 35-60 Gy were associated with increased stool frequency, where the strongest predictor was the 50 Gy dose-width with AUC 0.661 [95% CI 0.547, 0.776]. Results are shown in Figure 6.3d. The difference in 50 Gy dose-width between patients with or without increased stool frequency was 4.7 % (p = 0.009). Planned dose-widths were significant at 35-55 Gy. The planned 30 Gy dose-width produced a greater AUC than the corresponding accumulated measurement, but for all other significant dose-widths, accumulated dose produced a greater AUC. Analysis of 2D-DSMs found the accumulated 45 Gy isodose to be most predictive, with AUC = 0.648 [95% CI 0.536, 0.759].

Bowel bother $\geq$ G1 (n = 55, 39.6%)

The planned 45 Gy dose-width was most strongly associated with G1 bowel bother, with AUC = 0.625 [95% CI 0.525,0.725]. The difference in planned 45 Gy dose-width between patients with and without G1 bowel bother was 3.6 % (p = 0.026). Planned dose-widths from 35-55 Gy were all predictive of G1 bowel bother, as were accumulated dose-widths from 40-60 Gy and EUD. The 2D-DSM analysis found accumulated 60-Gy dose-width to be most predictive, with AUC = 0.638 [95% CI 0.543, 0.733].

Bowel bother $\geq$ G2 (n = 32, 23.0%)

Accumulated EUD was the only dosimetric parameter predictive of G2 bowel bother, with AUC = 0.615 [95% CI 0.503, 0.728]. However, the difference in accumulated EUD between patients with and without G2 bowel bother were not significant (p = 0.069). Planned dose-widths at 40 and 45 Gy, and accumulated dose-widths at 40 and 60 Gy produced AUCs greater than 0.6, but the lower bound of the 95%CI dropped below 0.5 so these could not be considered predictive of G2 bowel bother. The same result was found from 2D-DSM analysis, where accumulated EUD produced AUC = 0.648 [95% CI 0.536, 0.757].

Bowel bother $\geq$ G3 (n = 13, 9.4%)

Both planned and accumulated EUD, and the planned 45 Gy dose-width, produced AUCs greater than 0.6 for G3 bowel bother. The lower extent of the 95%CI fell below 0.5 so could not be considered significant. Despite not producing a significant result, the highest AUC was from the accumulated EUD (AUC = 0.650 [95% CI 0.496, 0.804]), which agreed with the 2D-DSM results.

6.2.3 Discussion

Dosimetric parameters extracted from both planned and accumulated FE-DSMs of the rectal wall were found to be associated with 9 of the 12 toxicity endpoints evaluated: diarrhoea, G1 and G2 proctitis, G1 and G2 rectal bleeding, RTOG G2 toxicity, frequency, G1 and G2 bowel bother. No significant associations were determined for faecal incontinence, RTOG G1 toxicity, or G3 bowel bother.

When comparing results between FE-DSM analysis and the 2D-DSM analysis conducted in Chapter 4, there is a combination of both complementary and conflicting data for the same patient datasets. Dosimetric parameters from FE-DSM were more predictive than those from 2D-DSM analysis for 6 of the 9 endpoints demonstrating dose-toxicity associations; G1 and G2 proctitis, G1 and G2 rectal bleeding, G2 RTOG toxicity, and frequency. 2D-DSM analysis produced greater AUCs for diarrhoea and G1 and G2 bowel bother, and also found associations between accumulated EUD and G1 RTOG toxicity and G3 bowel bother, which were not evident from FE-DSM analysis. Neither 2D-DSM nor FE-DSM analysis found any dosimetric parameter investigated to be significantly associated with faecal incontinence. An improvement in toxicity prediction was expected using the FE-DSM method due to the more robust approach to anatomical modelling. However, differences were relatively small, with an average improvement in AUC of 0.022 (range 0.4 - 3.9). In general, the results suggest a benefit to using FE-DSMs to represent dose to the rectal wall when evaluating dose-toxicity predictions using dose-widths and EUD.

Dosimetric parameters from accumulated dose was more discriminative of toxicity than planned dose for 5 of 9 endpoints demonstrating dose-toxicity associations; G1 and G2 rectal bleeding, RTOG G2 toxicity, G2 bowel bother, and frequency. Generally, dose-widths from 50-60 Gy produced the highest AUCs, ranging from 0.652 for G1 rectal bleeding to 0.805 for G2 rectal bleeding. Dosimetric parameters from planned dose produced stronger associations than accumulated delivered dose for diarrhoea, G1 and G2 proctitis, and G1 bowel bother. Across these four endpoints, dose-widths between 35-45 Gy consistently produced the highest AUCs, ranging from 0.615 for incontinence to 0.688 for G1 proctitis. Stronger associations were observed for dose-widths than EUD across 8 out of 9 endpoints with significant outcomes, suggesting that spatial dose features are more strongly associated with toxicity outcomes than a one-dimension dose descriptor, when evaluating both planned and delivered dose using FE-DSMs.

For diarrhoea, the planned 35 Gy dose-width was the strongest predictor, common to both 2D-DSM and FE-DSM analysis. For proctitis, FE-DSM analysis shows the planned 45 Gy dose-width to be associated with both G1 and G2 proctitis (equally predictive as the 40 Gy dose-width for the latter). For G2 proctitis, the strongest 2D-DSM correlation was observed

for the accumulated 60 Gy dose-width, and the planned 40 and 45 Gy dose-widths were not found to be significant. In the 2D-DSM analysis, although the planned 45 Gy dose-width was found to be associated with G1 proctitis, the planned 65 Gy dose-width produced the greatest AUC.

The 2D-DSM analysis found accumulated EUD to be most strongly associated with G1 rectal bleeding, whereas FE-DSM analysis found accumulated 55 Gy dose-width to be more predictive than accumulated EUD. However, results were quite similar and differences could not be considered statistically significant. For G2 rectal bleeding, 2D-DSM analysis found the planned 65 Gy dose-width to be most predictive, whereas the strongest AUC from FE-DSM was the accumulated 60 Gy dose-width. From FE-DSM analysis, the accumulated 60 Gy dose-width produced an AUC of over 0.8 which is considered highly significant, but was not a significant improvement from the AUC from the equivalent planned dose parameter.

Results from 2D-DSM and FE-DSM analysis were in agreement for G2 RTOG toxicity, both showing accumulated 50 Gy dose-width as the strongest predictor. Close agreement was observed for increased stool frequency, where the strongest predictors from 2D-DSM and FE-DSM analysis were accumulated 45 and 50 Gy dose-widths, respectively.

For bowel bother, 2D-DSM analysis found the accumulated 60 Gy dose-width to be most strongly associated with G1 symptoms. This was also a significant predictor in FE-DSM analysis, but the planned 45 Gy dose-width produced a higher AUC. For G2 bowel bother, accumulated EUD was found to be most discriminative from both 2D-DSMs and FE-DSMs.

Delivered accumulated dose to the rectal wall was different from planned dose for all patients. In general, accumulated dose was lower than planned dose when comparing using EUD. Significant predictors were found for 9 of 12 clinical endpoints examined. For 6 of these 9 endpoints, using FE-DSMs improved the strength of the dose-toxicity association. Of these, accumulated dose provided a stronger predictor toxicity than planned dose for 4 endpoints, although some differences were not significant. Comparing AUCs with previous 2D-DSM results (Chapter 4), FE-DSM findings supported some conclusions and conflicted with others. This may be because the difference in measured dose-width between 2D-DSM and FE-DSM increases with increasing dose level. FE-DSM revealed spatial dose parameters to be stronger predictors of toxicity than EUD for 8 out of 9 significant dose-toxicity associations. Improved correlations between spatial dose features and toxicity may be an indicator that using FE biomechanical modelling is a more accurate method for dose accumulation than the 2D expansion method described in Chapter 4.

However, the advantage of accumulated dose in discriminating between toxicity outcomes over planned dose was small. The magnitude of the differences, and widths of the 95% CIs were not significantly different at the highest AUCs. The dose parameters investigated here

were based on previous methods described in the literature [27], and may not be optimised for analysing delivered dose accumulated using FE modelling. Further methods for quantifying accumulated dose will be explored in the following section.

6.3 Voxel-based geometric subregion analysis

6.3.1 Materials and methods

The investigation in Section 6.2, compares the FE-DSM approach with the 2D-DSM approach and analyses spatial dose metrics (dose-widths) with a single metric describing a weighted mean dose across the entire rectal wall (EUD). Figure 6.4 shows the difference in FE-DSM between patients with and without rectal bleeding. The difference map shows a complex inhomogeneous dose distribution, where the regions showing the largest differences are located away from the central region where dose-widths were measured. For this reason, we explore spatial features of FE-DSMs beyond dose-widths.

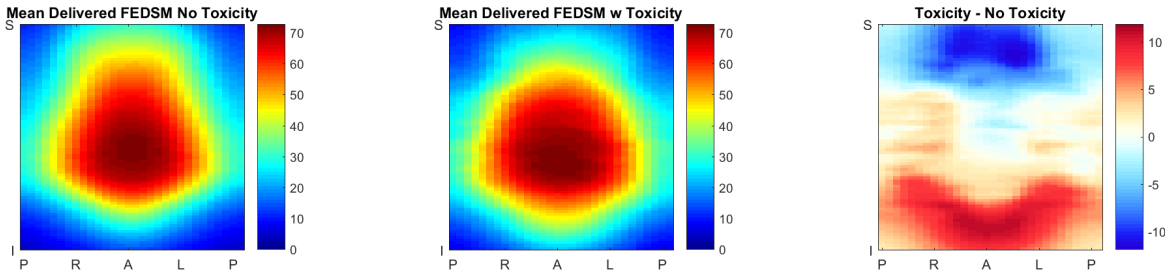


Fig. 6.4 Mean difference in accumulated delivered FE-DSM between patients with and without grade 2 rectal bleeding. S = superior, I = inferior, P = posterior, A = anterior, R = right, L = left.

From the VoxTox consolidation cohort, 139 prostate cancer patients were included in the analysis. Patient characteristics and toxicity endpoints included were presented previously in Section 6.2.1. Figure 6.5 shows how the different geometric subregions were defined on the FE-DSM, divided (a) longitudinally into anterior (ant) and posterior (post) hemispheres, (b) laterally into superior (sup), middle (mid), and inferior (inf) thirds, as well as (c) longitudinal quarters, anterior, posterior, right, and left.

Each patient's planned and accumulated FE-DSMs were divided into a total of 12 subregions. The EUD of each subregion was calculated and used as the input for receiver operator characteristic (ROC) curve analysis. The area under the ROC curve (AUC) of each subregion's EUD was compared with the EUD of the full rectal wall.

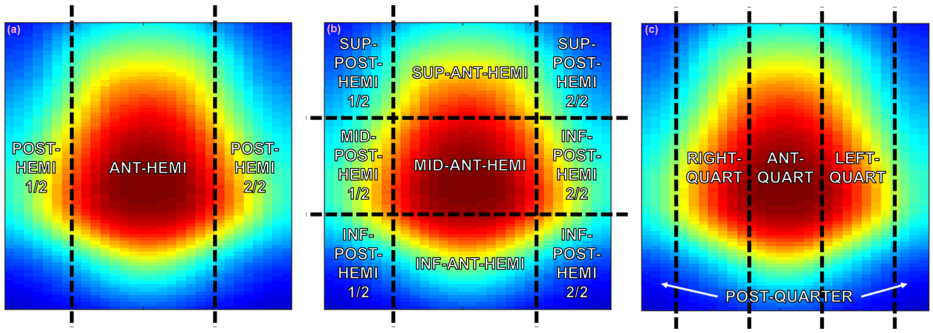


Fig. 6.5 Division of rectal DSMs into geometric subregions, SRR_{geom}. (a) Longitudinal hemi-rectum (anterior, ant, and posterior, post), (b) Longitudinal quarter-rectum (ant, post, left, right), (c) Ant and post hemi-rectum further split into lateral thirds (superior, sup; middle, mid; and inferior, inf).

6.3.2 Results

The strongest dose-toxicity association from a geometric subregion, from planned and accumulated FE-DSMs, are shown for all endpoints in Figure 6.6 and Table 6.3. Details of subregions found to be discriminative of toxicity are discussed per endpoint below.

Diarrhoea (n = 36, 25.9%)

The AUC of the accumulated EUD for the full FE-DSM of the rectal wall was 0.617 [95% CI 0.504, 0.731]. Planned EUD was not found to be significant. From the geometric subregions, the regions producing AUCs greater than 0.617 were: accumulated posterior hemi-rectum (AUC = 0.620 [95% CI 0.509, 0.731]), accumulated mid-third posterior hemi-rectum (AUC 0.623 [95% CI 0.512, 0.733]), accumulated and planned longitudinal post quarter (AUC 0.622 [0.513, 0.731], 0.645 [0.540, 0.750], respectively). The strongest predictor was found to be the accumulated longitudinal right quarter (AUC 0.647 [95% CI 0.535, 0.759]).

Faecal incontinence (n = 24, 17.3%)

No significant associations were observed for planned or accumulated EUDs of the full rectal wall, or from any geometric subregion.

Proctitis $\geq$ G1 (n = 26, 18.7%)

The AUC of the full rectal wall from the planned FE-DSM was 0.641 [95% CI 0.532, 0.751]. Accumulated EUD was not found to be predictive of G1 proctitis. The subregion with the greatest AUC was the planned posterior-hemi rectum (AUC 0.682 [95% CI 0.571, 0.794]). Other subregions with AUC greater than 0.641 were: planned longitudinal right and posterior quarters (AUC 0.678 [95% CI 0.575, 0.782], 0.660 [95% CI 0.535, 0.785], respectively),

accumulated longitudinal right quarter (AUC 0.646 [95% CI 0.528, 0.764]), and planned mid posterior hemi rectum (AUC 0.656 [95% CI 0.537, 0.774]).

Proctitis $\geq$ G2 (n = 20, 14.4%)

Planned and accumulated EUDs of the full rectal wall did not produce AUC discriminative of G2 proctitis. By splitting the FE-DSMs into subregions, significant associations were found for both planned and accumulated posterior hemi-rectum (AUC 0.647 [95% CI 0.525, 0.769], 0.640 [95% CI 0.508, 0.772], respectively), and both planned and accumulated longitudinal right quarters (AUC 0.648 [95% CI 0.533, 0.763], 0.662 [95% CI 0.535, 0.790], respectively).

Rectal bleeding $\geq$ G1 (n = 46, 33.1%)

For the full rectal wall FE-DSM, the accumulated EUD was 0.641 [95% CI 0.545, 0.737] and planned EUD was 0.615 [95% CI 0.520, 0.711]. The rectal subregions greater than or equal to 0.641 were: accumulated anterior and posterior hemis (AUC 0.641 [95% CI 0.545, 0.737], 0.642 [95% CI 0.545, 0.739], respectively), and planned and accumulated inferior-third posterior hemi-rectum (AUC 0.645 [95% CI0.544, 0.747], 0.654 [95% CI0.553, 0.755], respectively).

Rectal bleeding $\geq$ G2 (n = 16, 11.5%)

Accumulated EUD of the full rectal wall produced an AUC of 0.690 [95% CI 0.568, 0.812], and the AUC for planned EUD was 0.629 [95% CI 0.515, 0.742]. The greatest AUCs from subregions were for accumulated and planned posterior hemi-rectum (AUC 0.800 [95% CI 0.707, 0.894], 0.772 [95% CI 0.675, 0.869], respectively). Other subregions associated with AUCs greater than 0.690 were: planned and accumulated longitudinal right quarters (AUC 0.743 [95% CI 0.651, 0.836], 0.759 [95% CI 0.645, 0.872]), planned and accumulated longitudinal posterior quarters (AUC 0.746 [95% CI 0.625, 0.868], 0.768 [95% CI 0.649, 0.886], respectively), and accumulated mid and inferior posterior hemi-rectum (AUC 0.697 [95% CI 0.563, 0.831], 0.693 [95% CI 0.543, 0.844], respectively).

RTOG GI toxicity $\geq$ G1 (n = 73, 52.5%)

EUDs of the full rectal FE-DSM were not predictive of G1 RTOG toxicity. Splitting the rectal wall into geometric subregions revealed an association with accumulated longitudinal right quarter (AUC 0.601 [95% CI 0.507, 0.695]), but no other planned or accumulated dose metric was significant.

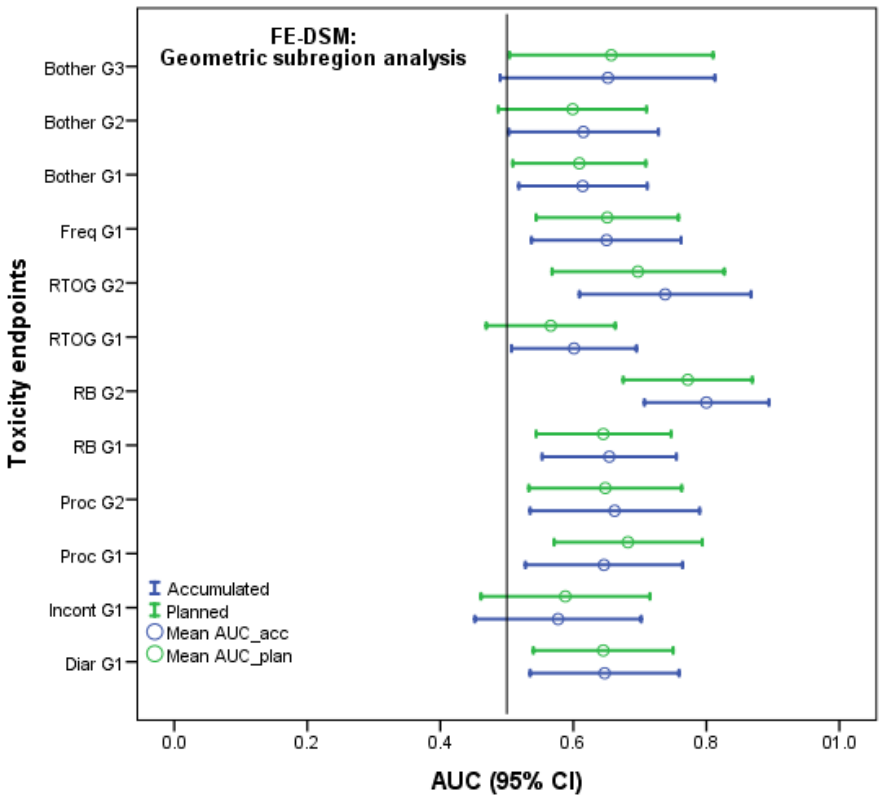


Fig. 6.6 Area under the receiver operator characteristic curve (AUC) and 95% confidence intervals (CI) for planned (plan) and accumulated (acc) equivalent uniform dose calculated from geometrically defined subregions for 12 toxicity endpoints. For each endpoint, the subregion demonstrating the strongest association has been plotted. Results were considered significant where mean AUC ≥ 0.6 and lower 95% CI ≥ 0.5.

RTOG GI toxicity $\geq$ G2 (n = 22, 15.8%)

For G2 RTOG toxicity, planned and accumulated EUDs of the full rectal FE-DSMs were not considered significant. When dividing into geometric SRRs, accumulated and planned longitudinal posterior quarter were most strongly associated with G2 RTOG GI toxicity, with AUCs of 0.738 [95% CI 0.609, 0.867], 0.697 [95% CI 0.568, 0.827], respectively. Significant associations were also observed for accumulated and planned posterior hemi-rectum (AUC 0.731 [95% CI 0.601, 0.861], 0.689 [95% CI 0.566, 0.813], respectively), accumulated and planned longitudinal right quarter (AUC 0.701 [95% CI 0.566, 0.836], 0.667 [95% CI 0.544, 0.791], respectively), and accumulated and planned mid posterior hemi-rectum (AUC 0.696 [0.562, 0.831], 0.639 [0.511, 0.768], respectively).

Stool frequency (n = 34, 24.5%)

The subregions found to be associated with increased stool frequency were: planned and accumulated posterior quarter (AUC 0.651 [95% CI 0.544, 0.758], 0.632 [95% CI 0.520, 0.744], respectively), accumulated and planned posterior hemi-rectum (0.650 [95% CI 0.537, 0.762], 0.619 [95% CI 0.505, 0.733], respectively), accumulated longitudinal right quarter (AUC 0.646 [95% CI 0.531, 0.760]), and accumulated mid posterior hemi-rectum (AUC 0.631 [95% CI 0.515, 0.747]). Neither planned nor accumulated EUD of the full rectal wall were associated with stool frequency.

Bowel bother $\geq$ G1 (n = 55, 39.6%)

Accumulated EUD for the full rectal FE-DSM was predictive of G1 bowel bother, with AUC 0.614 [95% CI 0.517, 0.710]. This AUC was found to be matched by the accumulated EUD of the longitudinal right quarter subregion (AUC 0.614 [95% CI 0.518, 0.711]). Planned posterior quarter was also also associated with G1 bowel bother, with AUC of 0.609 [95% CI 0.509, 0.709].

Bowel bother $\geq$ G2 (n = 32, 23.0%)

The accumulated EUD of the full rectal wall was found to be associated with G2 bowel bother (AUC 0.615 [95% CI 0.503, 0.728]). Accumulated anterior hemi-rectum and longitudinal right quarter (AUC 0.615 [95% CI 0.501, 0.728], 0.615 [95% CI 0.502, 0.729], respectively) were found to be equally predictive. No planned FE-DSM parameter was significantly associated with G2 bowel bother.

Bowel bother $\geq$ G3 (n = 13, 9.4%)

The only FE-DSM rectal subregion associated with G3 bowel bother was the planned longitudinal right quarter, with AUC 0.657 [95% CI 0.504, 0.810]. No planned FE-DSM parameter was significantly associated with G3 bowel bother.

Table 6.3 Geometric rectal subregions at risk (SRR_{geom}) most discriminative of toxicity from planned and accumulated dose surface maps generated using finite element modelling. SRR_{geom} producing the greatest area under the receiver operator characteristic curve (AUC) are listed. Endpoint entries were omitted where no parameter from planned or accumulated dose produced an AUC with minimum 95% confidence interval greater than 0.5, so was not considered significant. A dashed entry indicates that no significant parameter was determined. Bold indicates greatest AUC between planned and accumulated results.

	Planned dose		Accumulated dose	
Toxicity endpoint	SRR_{geom}	AUC	SRR_{geom}	AUC
Diarrhoea	post-quarter	0.645	right-quarter	**0.647**
Proctitis $\geq$ G1	post-hemi	**0.682**	right-quarter	0.646
Proctitis $\geq$ G2	right-quarter	0.648	right-quarter	**0.662**
Rectal bleeding $\geq$ G1	inf-post-hemi	0.645	inf-post-hemi	**0.654**
Rectal bleeding $\geq$ G2	post-hemi	0.772	post-hemi	**0.800**
GI toxicity $\geq$ G1	-	-	right-quarter	**0.601**
GI toxicity $\geq$ G2	post-quarter	0.697	post-quarter	**0.738**
Stool frequency	post-quarter	**0.651**	post-hemi	0.650
Bowel bother $\geq$ G1	post-quarter	0.609	EUD/right-quarter	**0.614**
Bowel bother $\geq$ G2	-	-	EUD/ant-hemi/right-quarter	**0.615**
Bowel bother $\geq$ G3	-	-	right-quarter	**0.657**

6.3.3 Discussion

For all endpoints where significant dose-toxicity associations were determined, the AUCs from geometric subregions were greater than those for the full rectal wall - with the exception of bowel bother greater than grade 1 and grade 2 where the effects were equal. No association was detected for faecal incontinence. For all other endpoints, the strongest AUCs were found from a range of different geometric subregions. There were 5 cases for accumulated dose, and 4 from planned dose, where the EUD of the full rectal wall was not found to be associated with the toxicity endpoint, but upon further division into geometrical subregions, significant associations were uncovered.

For 11 of 12 toxicity endpoints with significant dose-toxicity associations, the strongest association was found from accumulated dose. Subregions located posteriorly (n=6) and to the patient's right (n=4) revealed the strongest associations, with the exception of G2 bowel bother which found the longitudinal right-quarter and anterior-hemi to be equally predictive.

For 9 of the 11 toxicity endpoints with significant dose-toxicity associations, significant AUCs were found for a geometric subregion on the right side of the rectal wall. It is expected that dose to the right and left sides of the rectum should be approximately equal, so this effect may be due to an unknown clinical or anatomical phenomenon. This result may suggest that the rectum is able to expand more to the right, where differences were observed between planned and accumulated dose, which may be due to the continuation at the rectosigmoid junction where the rectum moves anteriorly and to the left.

EUD calculated to geometric subregions of the rectal wall demonstrated stronger associations with toxicity than EUD calculated to the full rectal FE-DSM. This result supports the hypothesis that the rectum has non-homogeneous radio sensitivity. Dose to the posterior and right side of the rectal wall were most discriminative of toxicity, suggesting that low and intermediate doses are more associated with toxicity than the high dose region. This effect has previously been reported by other groups using planned dose data [76, 87, 97, 115]. Furthermore, results indicate that different regions of the rectal wall may be associated with different toxicity endpoints, suggesting that the pathophysiology of toxicity side-effects differs per endpoint [70].

Overall, accumulated dose was more predictive of toxicity than planned dose, and dose to geometric subregions revealed stronger associations than dose to the full rectal wall. Knowledge of specific subregions associated with toxicity could be considered as regions to avoid during the treatment planning process, and as areas to monitor when concerns arise with regards to adaptive radiotherapy, to prospectively minimise the risk of developing toxicity.

6.4 Voxel-based statistical subregion analysis

6.4.1 Materials and methods

For each of the 12 toxicity endpoints considered (Table 6.1), patients in the training set were split according to those with toxicity, and those without. The FE-DSMs of prostate radiotherapy patients prescribed 60 Gy in 20 fractions were radiobiologically converted to the equivalent dose in 37 fractions using an α/β ratio of 2.149 (as described in Section 3.3). Planned and accumulated FE-DSMs were compared. It was postulated that specific regions of the rectal wall were more radiosensitive and hence more strongly associated with toxicity. The purpose of this investigation was to determine the areas of increased radiosensitivity associated with different toxicity endpoints to interrogate the dose differences specific to these regions.

For each voxel of the 30×80 FE-DSM, a two-sample t-test for equality of means was performed in MATLAB® (Mathworks®, Natick, MA). The two groups were defined by patients with and without toxicity and the null hypothesis assumed no difference in voxel dose between groups. Voxels were identified in areas where the dose was significantly higher for patients with toxicity than those without. Test were set to reject the null hypothesis at the 5 % significance level ($p < 0.05$).

Where the p-value at each voxel was less than the significance level, the dose difference between patients with and without toxicity was considered significant, and the null hypothesis was rejected. From the map of voxel-wise p-values (example shown for grade 2 rectal bleeding in Figure 6.7), a binary mask was generated where a value of 1 was assigned when the null hypothesis was rejected, and a value of 0 was assigned otherwise (Figure 6.8). Post-processing performed on the resulting binary masks included removing clusters of fewer than 10 pixels (unless the cluster would have been connected to a larger cluster had it not been for the posterior splitting of the FE-DSM for visualisation in 2D), applying a least-squares smoothing filter (Savitzky-Golay filter in the MATLAB image processing toolbox), and filling in any remaining holes. The function of the post-processing stage was to smooth sharp edges and remove noise. This identified the rectal subregion at risk corresponding to a p-value < 0.05, SRR_{pmap}.

The issue of correcting for multiple comparison testing in voxelwise dose-toxicity studies has been debated in the scientific literature. Here, the parameter of interest for differentiating between patients with and without toxicity is the spatial pattern of the resulting distribution, rather than the significance level of individual pixels, so conventional approaches to multiple testing may be unnecessarily restrictive and overly conservative. In their analysis of bladder dose-surface maps, Palorini et al. [121] discussed the issue at depth, and argued against

adjusting p-values in this context, emphasising that their main goal was 'to identify the shape of dose regions that show the highest discrimination between the two groups of patients (toxicity versus no toxicity) without looking at significance at the single-pixel level'. They applied a similar post-processing smoothing approach and removed small regions with $p < 0.05$. Based on these recommendations, Onjukka et al.[115] reported results for unadjusted p-values in a recent pattern analysis of rectal dose surface maps. Data for adjusted p-values were also included for completeness. Their results support the concept that adjusting the p-values has the effect of being increasing sensitivity to the point of over-fitting to the data when applied to voxel-level DSM analysis. Because of this paradigm shift towards using unadjusted p-values for evaluating the spatial patterns of DSMs between patients with and without toxicity, the unadjusted p-values have been applied in this analysis. However, to complement the methodology applied by Dréan et al. [61] in their work on subregion analysis, their selected approach to multiple testing using the Benjamini-Hochberg procedure [17, 69, 146] was also investigated.

The Benjamini-Hochberg procedure involves determining a p-value threshold by ordering and ranking voxels based on their p-value test statistic. The Benjamini-Hochberg critical value was calculated using the formula $(j/m)q^*$, where j is the rank number assigned to the voxel's p-value, m is the total number of tests (= 2400), and q^* is the false discovery rate (taken to be 0.2 [121]). The test with the highest rank, for which the p-value is less than the Benjamini-Hochberg critical value, determines the final modified p-value threshold (replacing the previous 0.05 value).

Voxel-level analysis was used to determine larger spatial patterns of interest. The final processed binary masks derived using unadjusted and adjusted p-values represent the probabilistic rectal subregions at risk. The binary masks shown in Figures 6.8 and 6.9 were applied

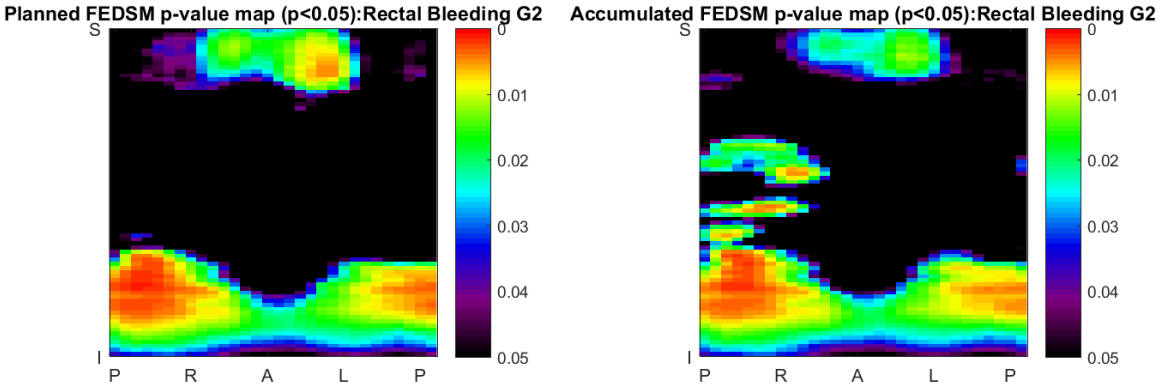


Fig. 6.7 Voxel-wise probability map for 139 patients with (n = 16) and without (n = 123) grade 2 rectal bleeding, for planned and accumulated dose. Test statistics with $p < 0.05$ are displayed.

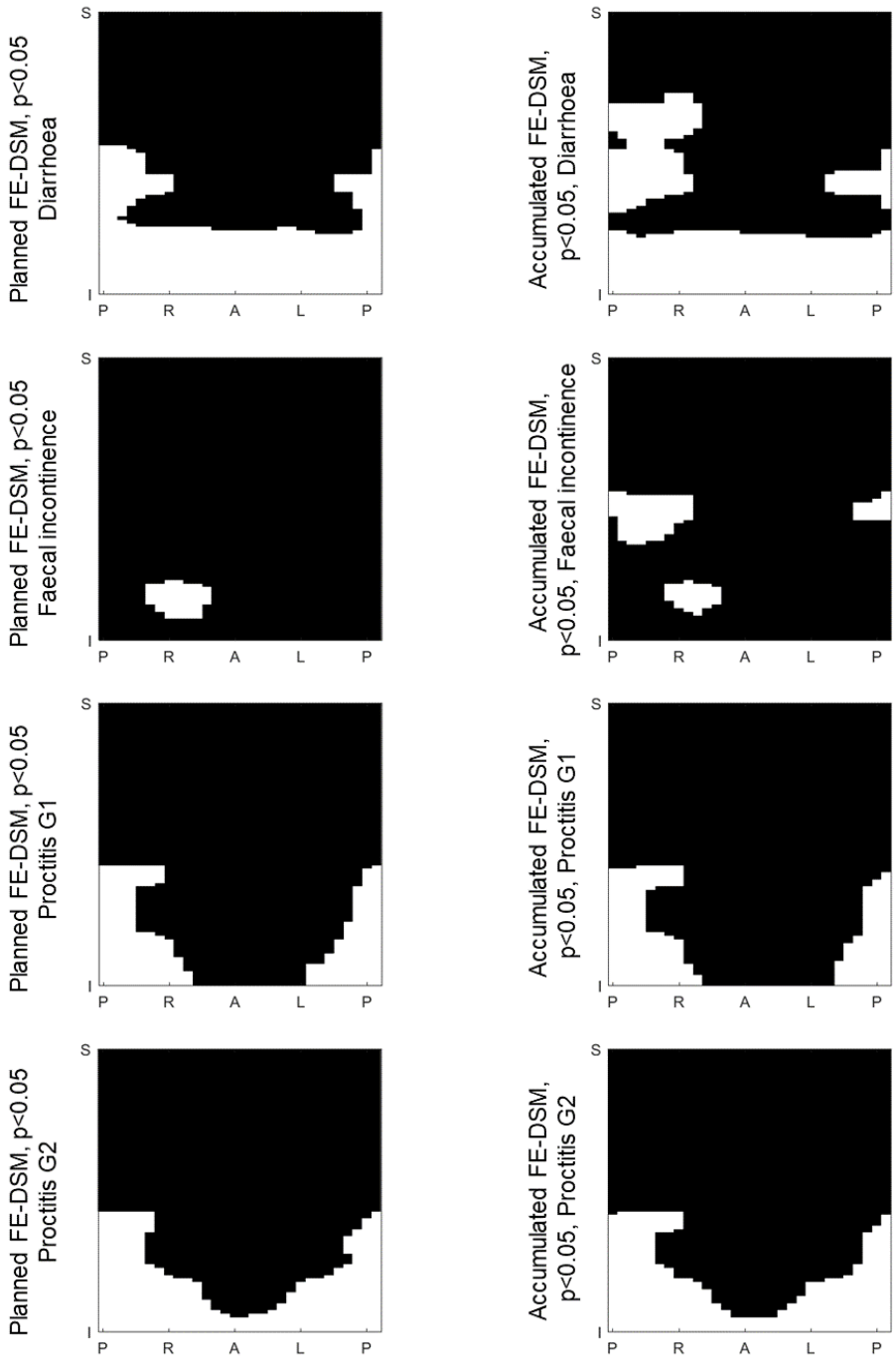
to each patient's planned and accumulated FE-DSMs such that only the dose to voxels from the non-zero (white) regions were considered. The EUD was then calculated based on the voxel dose within these subregions. ROC analysis was performed for the endpoints that produced a binary mask, and AUCs were reported if greater than 0.6 with a minimum 95%CI greater than 0.5 [71].

6.4.2 Results

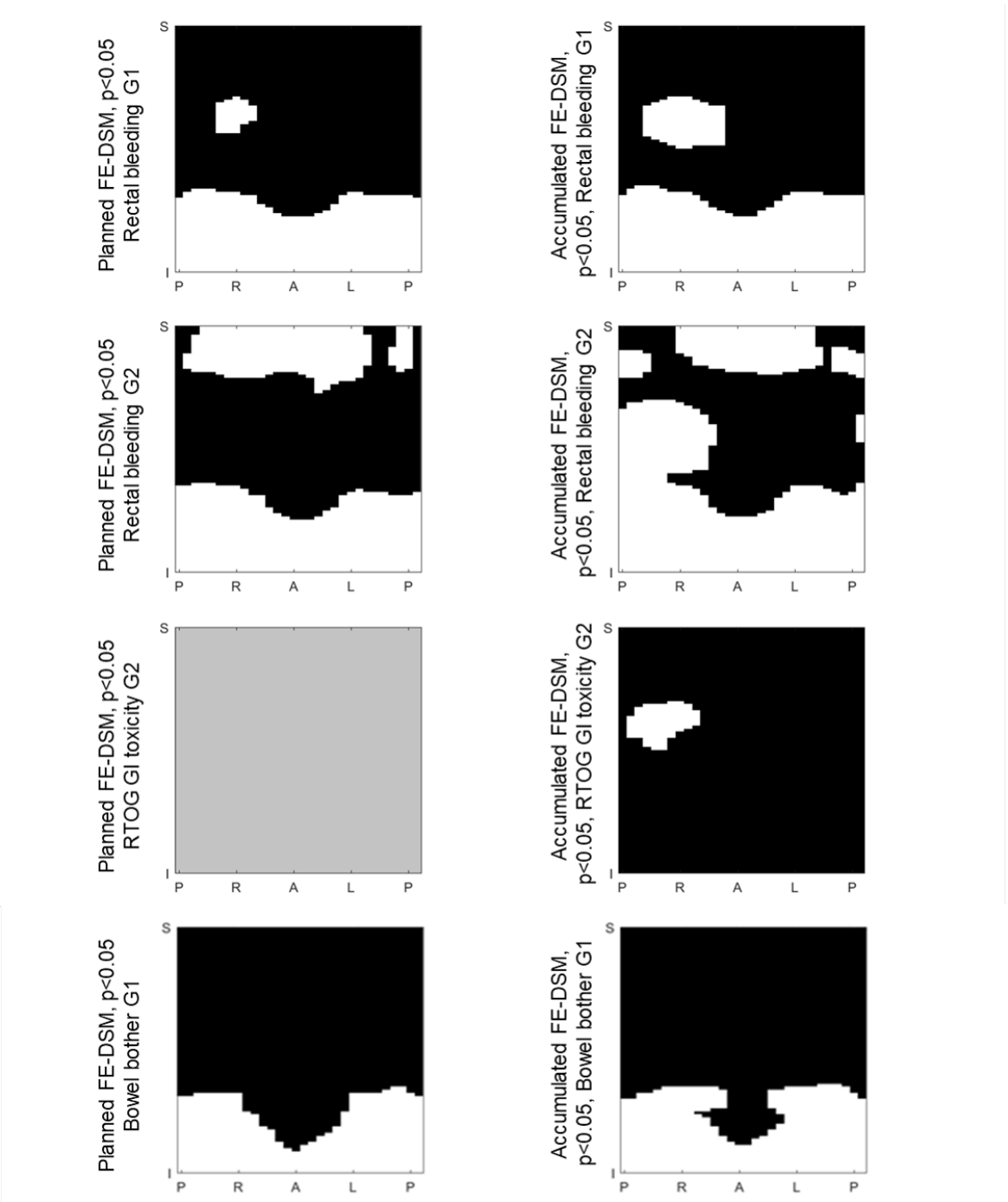
The binary masks generated using the unadjusted p-value of 0.05 are shown in Figure 6.8 for 8 toxicity endpoints where areas of significant dose patterns were identified between patients with and without toxicity. These correspond to the SRR_{pmap} that resulted in significant associations with the toxicity endpoint of interest, and are discussed below. The only endpoint not to produce a binary mask indicating spatial regions of significant dose difference was the planned FE-DSM for RTOG G2 GI toxicity. Binary masks using the Benjamini-Hochberg adjusted p-value were generated for 6 endpoints and have been included for comparison in Figure 6.9. Where binary masks could not be generated indicates that no areas of significant dose difference could be identified. For the endpoints where probabilistic subregions were identified, the binary mask was applied to each patient's FE-DSM individually. The EUD was calculated based on the voxels in the probabilistic subregion, and ROC analysis was performed. Results for unadjusted p-value binary masks are presented in Table 6.4 discussed below where SRR_{pmap} revealed an association with toxicity, with AUC ≥ 0.6 and lower 95% CI ≥ 0.5.

Diarrhoea (n = 36, 25.7%)

The binary mask for $p < 0.05$ showed significant differences for patients with and without diarrhoea in a band across the inferior quarter, with a mid-posterior cluster for both planned and accumulated FE-DSMs. The mid-posterior cluster appeared more prominently towards the right for accumulated dose. Accumulated SRR_{pmap}, AUC 0.665 [95% CI 0.560, 0.770], was more strongly associated with diarrhoea than planned SRR_{pmap}, AUC 0.656 [95% CI 0.0.555, 0.757].



(a) Binary masks showing probabilistic subregions of FE-DSMs where voxel dose difference were significant ($p<0.05$) between patients with and without toxicity for Diarrhoea, faecal incontinence, proctitis $\geq$ G1, proctitis $\geq$ G2 (top to bottom).



(b) Binary masks showing probabilistic subregions of FE-DSMs where voxel dose difference were significant ($p<0.05$), as indicated in white, between patients with and without toxicity for rectal bleeding $\geq$ G1 and $\geq$ G2, RTOG gastrointestinal (GI) toxicity $\geq$ G2, and bowel bother $\geq$ G1 (top to bottom). Grey indicates no significant regions were identified.

Fig. 6.8

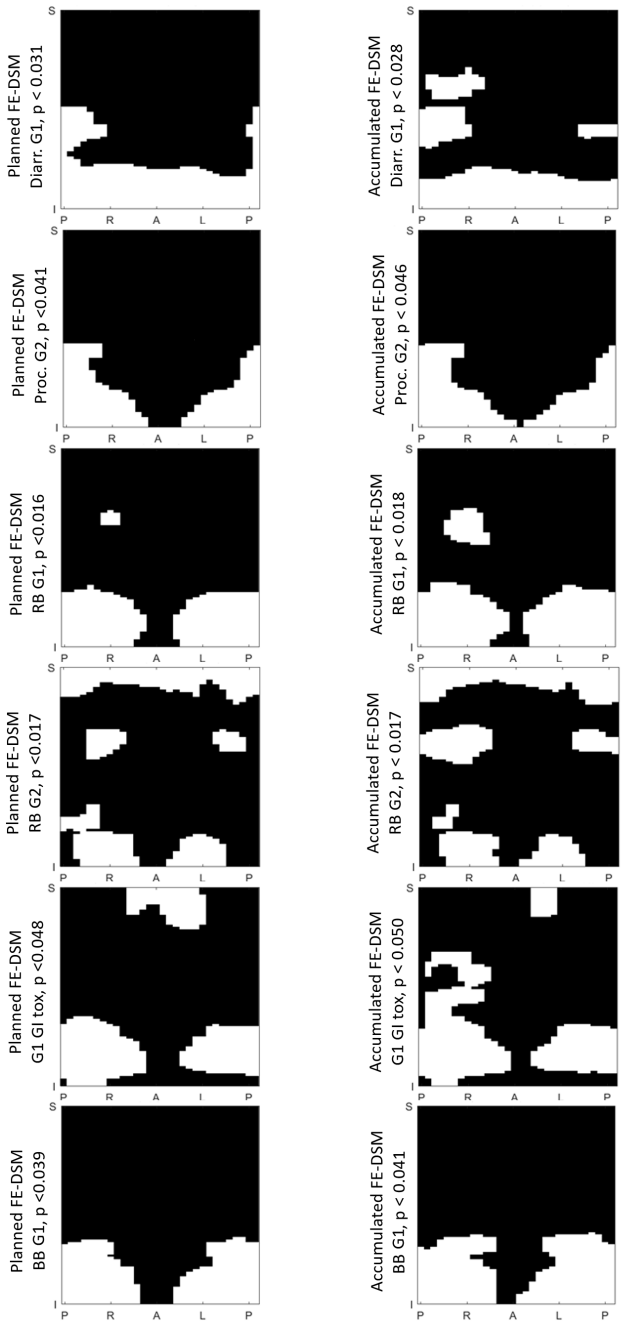


Fig. 6.9 Binary masks showing probabilistic subregions of FE-DSMs where voxel dose difference were significant between patients with and without toxicity, generated using Benjamini-Hochberg adjusted p-values. Binary masks were determined for 6 of 12 toxicity endpoints investigated: grade 1 diarrhoea (diar. G1), grade 2 proctitis (proc. G2), G1 & G2 rectal bleeding (RB), G1 gastrointestinal toxicity (GI tox), grade 1 bowel bother (BB G1).

Faecal incontinence (n = 24, 17.3%)

A cluster at the inferior right-lateral was common to both planned and accumulated FE-DSMs. A mid-posterior to mid-right cluster was also apparent on the accumulated $p < 0.05$ binary mask. Accumulated dose to SRR_{pmap}, AUC 0.633 [95% CI 0.506, 0.760] was associated with faecal incontinence, and planned dose was not considered significant.

Proctitis $\geq$ G1 (n = 26, 18.7%)

Significant clusters were observed at the 5 % significance level extending from the mid to inferior post-hemi rectal wall, common to planned and accumulated FE-DSMs. Planned dose to SRR_{pmap} produced the strongest association, with AUC 0.677 [95% CI 0.557, 0.797], compared to accumulated SRR_{pmap}, AUC 0.647 [95% CI 0.525, 0.770].

Proctitis $\geq$ G2 (n = 20, 14.4%)

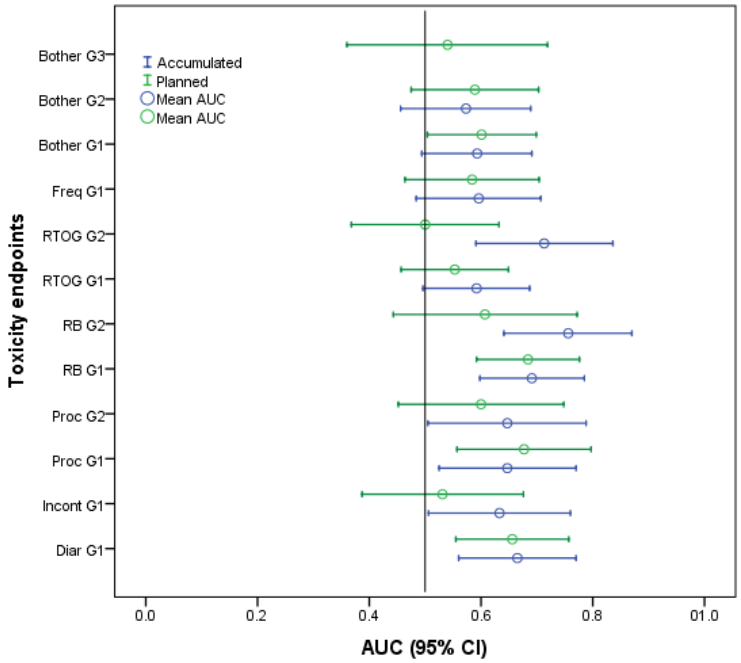
Similar $p < 0.05$ cluster patterns were observed for G2 proctitis as described for G1 proctitis, although compared to the previous cluster pattern, clusters were connected by a band at the inferior end of the rectal wall, spanning the full circumference. The accumulated SRR_{pmap} was associated with G2 proctitis, with AUC 0.647 [95% CI 0.505, 0.788], and no significant association was found for planned SRR_{pmap}.

Rectal bleeding $\geq$ G1 (n = 46, 33.1%)

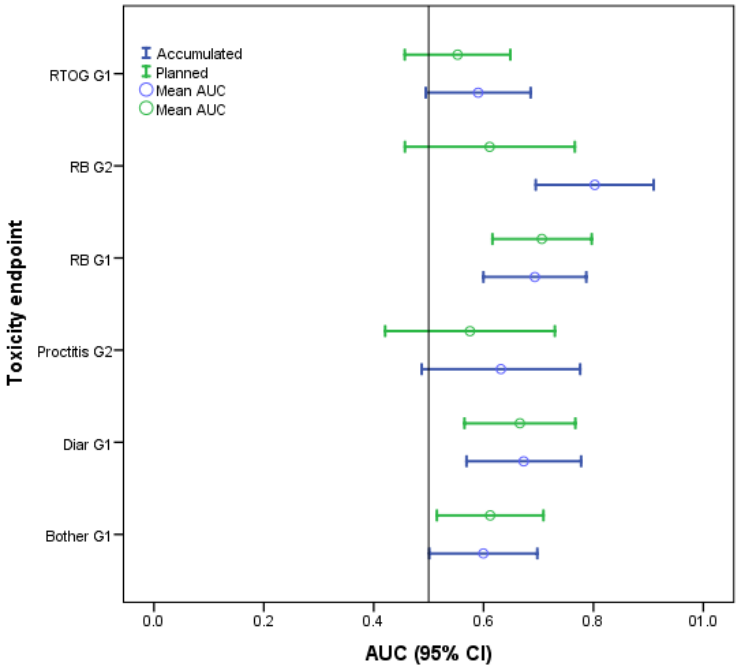
A wide band spanning right to left posteriorly at the inferior third of the rectal wall was common to both planned and accumulated SRR_{pmap} for G1 rectal bleeding, as was a smaller cluster in the mid-right lateral region which was larger for accumulated dose than planned dose. Accumulated dose (AUC 0.691 [95% CI 0.598, 0.785]) was more strongly associated with G1 rectal bleeding than planned dose (AUC 0.684 [95% CI 0.592, 0.776]).

Rectal bleeding $\geq$ G2 (n = 16, 11.5%)

Larger regions of significant dose differences were observed for G2 rectal bleeding than for G1. Clusters spanned the circumference of the inferior third, and most of the superior third circumference for planned and accumulated dose, with additional expansion into the mid-posterior-right lateral region on the accumulated binary mask. Accumulated dose to SRR_{pmap} was most strongly discriminative of G2 rectal bleeding, AUC 0.756 [95% CI 0.641, 0.870], than planned SRR_{pmap}, which was not considered significant due to the relatively large width of the 95% CI.



(a) FE-DSM SRR_{pmap} AUCs from unadjusted p-values ($p \leq 0.05$)



(b) FE-DSM SRR_{pmap} AUCs from adjusted p-values

Fig. 6.10 Area under the receiver operator characteristic curve (AUC) and 95% confidence intervals (CI) for planned (plan) and accumulated (acc) equivalent uniform dose (EUD) calculated from rectal subregions at risk (SRR) defined on FE-DSMs using unadjusted and adjusted p-values. Twelve toxicity endpoints were analysed; results are shown only where a binary mask of regions of significant dose difference between patients with and without toxicity could be generated.

Table 6.4 Area under the receiver operator characteristic curve (AUC) indicating discriminative ability of probabilistic rectal subregions at risk EUD, for different toxicity endpoints, calculated using both planned and accumulated finite element generated dose surface maps. Dash entries indicate no binary mask was generated, or result was not significant.

Toxicity endpoint	AUC - Planned dose	AUC - Accumulated dose
Diarrhoea	0.656	**0.665**
Incontinence	-	**0.633**
Proctitis $\geq$ G1	**0.677**	0.647
Proctitis $\geq$ G2	0.600	**0.647**
Rectal bleeding $\geq$ G1	0.684	**0.691**
Rectal bleeding $\geq$ G2	0.607	**0.756**
GI toxicity $\geq$ G2	-	**0.713**
Bowel bother $\geq$ G1	**0.601**	-

RTOG GI toxicity $\geq$ G2 (n = 22, 15.8%)

Accumulated SRR_{pmap} revealed a single cluster at the mid-right-posterior area of the rectal wall. No binary mask could be generated using planned FE-DSMs. Accumulated SRR_{pmap} resulted in an AUC of 0.713 [95% CI 0.591, 0.836].

Bowel bother $\geq$ G1 (n = 54, 38.8%)

For both planned and accumulated binary masks, clusters spanned the circumference of the inferior rectal wall, dipping inferiorly at the most anterior point. Planned SRR_{pmap} was associated with G1 bowel bother, with AUC = 0.601 [95% CI 0.504, 0.699], and accumulated SRR_{pmap} was not found to be significant.

6.4.3 Discussion

Subregions of the rectal wall were statistically identified by interrogating dose differences at the voxel level, comparing FE-DSMs of patients with and without toxicity. Several binary masks found clusters in the inferior part of the rectum to be significant for different toxicities, and a few indicated regions of significant dose difference on the right side of the rectal wall. This depth of voxel-based toxicity analysis has not previously been reported for accumulated dose to the rectum. Clusters of pixels were identified that facilitated higher resolution interrogation of associations between rectal subregions and toxicity, beyond the previous geometric-based approach described in Section 6.3. Results complement the results of the geometric subregion analysis by supporting the hypothesis that the rectum has non-homogeneous radiosensitivity [3, 87, 141], and that particular areas are more strongly

associated with different toxicity endpoints. This effect has also been reported in the literature for planned dose to the rectal wall [34, 115].

EUD of SRR_{pmap} revealed significant dose-toxicity correlations for 8 of 12 endpoints. Accumulated dose produced stronger associations than planned dose for 6 of the 8 toxicity endpoints; diarrhoea, faecal incontinence, G2 proctitis, G1 rectal bleeding, and RTOG G2 GI toxicity. However, for endpoints other than G2 rectal bleeding and RTOG G2 GI toxicity, differences between accumulated dose and planned dose were small. Planned dose difference was found not to be significant at $p < 0.05$ between patients with and without RTOG G2 GI toxicity so no binary mask could be created.

Results were calculated using an unadjusted p-value of 0.05. To address the issue of multiple comparisons, results were also calculated based on adjusted p-values, as shown in Figure 6.10b. Binary masks could only be generated for 6 toxicity endpoints based on adjusted p-values (shown in Figure 6.9). Where comparative measurements were available, similar trends were observed between adjusted and unadjusted results, and notably the accumulated AUC became significantly greater than planned AUC for G2 rectal bleeding when using the adjusted p-value. However, this study focused on the use of the unadjusted p-value because, as argued by Palorini et al. [121], it is the shape of the identified subregion that is of interest, not the test statistic of each individual pixel. This way, maximum information is being extracted in order to identify regions of the rectum most associated with rectal toxicity.

6.5 Concluding discussion

6.5.1 Comparing analysis and parameterisation approaches

In this chapter, FE-DSM analysis is compared with 2D-DSM analysis using the same dose parameterisation approach (dose-widths and EUD). Increasingly complex methods of dose parameterisation were then investigated, from spatially determined geometric subregions (SRR_{geom}), to subregions identified using statistically derived probability maps (SRR_{pmap}) based on unadjusted and adjusted p-values. Here, we compare the results presented for each analysis and parameterisation approach presented.

For planned and accumulated dose, the average AUC of the parameter found to be most discriminative of each toxicity endpoint, per analysis/parameterisation approach: 2D-DSM dose-width/EUD, FE-DSM dose-width/EUD, FE-DSM SRR_{geom} region, and FE-DSM SRR_{pmap} (unadjusted), are presented in Figures 6.11 and 6.12.

For planned dose, 9 of 12 endpoints resulted in significant associations between dose and toxicity. Of these, FE-DSM analysis (dose-width/EUD, SRR_{geom}, or SRR_{pmap}) produced greater AUCs than 2D-DSM analysis for 8 of 9 endpoints. Parameterising FE-DSMs using dose-widths resulted in larger AUCs than EUD for 5 of these 8 endpoints. SRR_{geom} produced the greatest AUCs for 2 endpoints, and SRR_{pmap} for 1 endpoint.

For accumulated dose, all 12 endpoints resulted in significant associations between dose and toxicity. Of these, the largest AUC was found using FE-DSM analysis (dose-width/EUD, SRR_{geom}, or SRR_{pmap}) for 8 endpoints. The 4 endpoints where 2D-DSM analysis resulted in the largest AUC were for the 3 different grades of bowel bother and G1 RTOG GI toxicity. For FE-DSM analysis, SRR_{pmap} resulted in the greatest AUC for 4 endpoints, SRR_{geom} for 1, and dose-widths for 2 endpoints. The general trend across individual endpoints with significant dose-toxicity associations was that AUC improved with increasing complexity of parameterisation technique.

Overall, results support the hypothesis that using FE-DSMs to analyse dose to the rectal wall improves the discriminative power of dose-toxicity predictions when compared to the 2D-DSM approach. However, this was not consistent across all endpoints examined, and the differences were generally not statistically significant. A direct comparison between 2D-DSM and FE-DSM using dose-widths and EUD found results to be significantly different. Where stronger dose-toxicity associations were observed with increasing complexity of dose parameterisation, this was generally more pronounced for accumulated dose than planned dose. Results are encouraging and further research into advanced parameterisation approaches tailored to FE-DSMs is recommended.

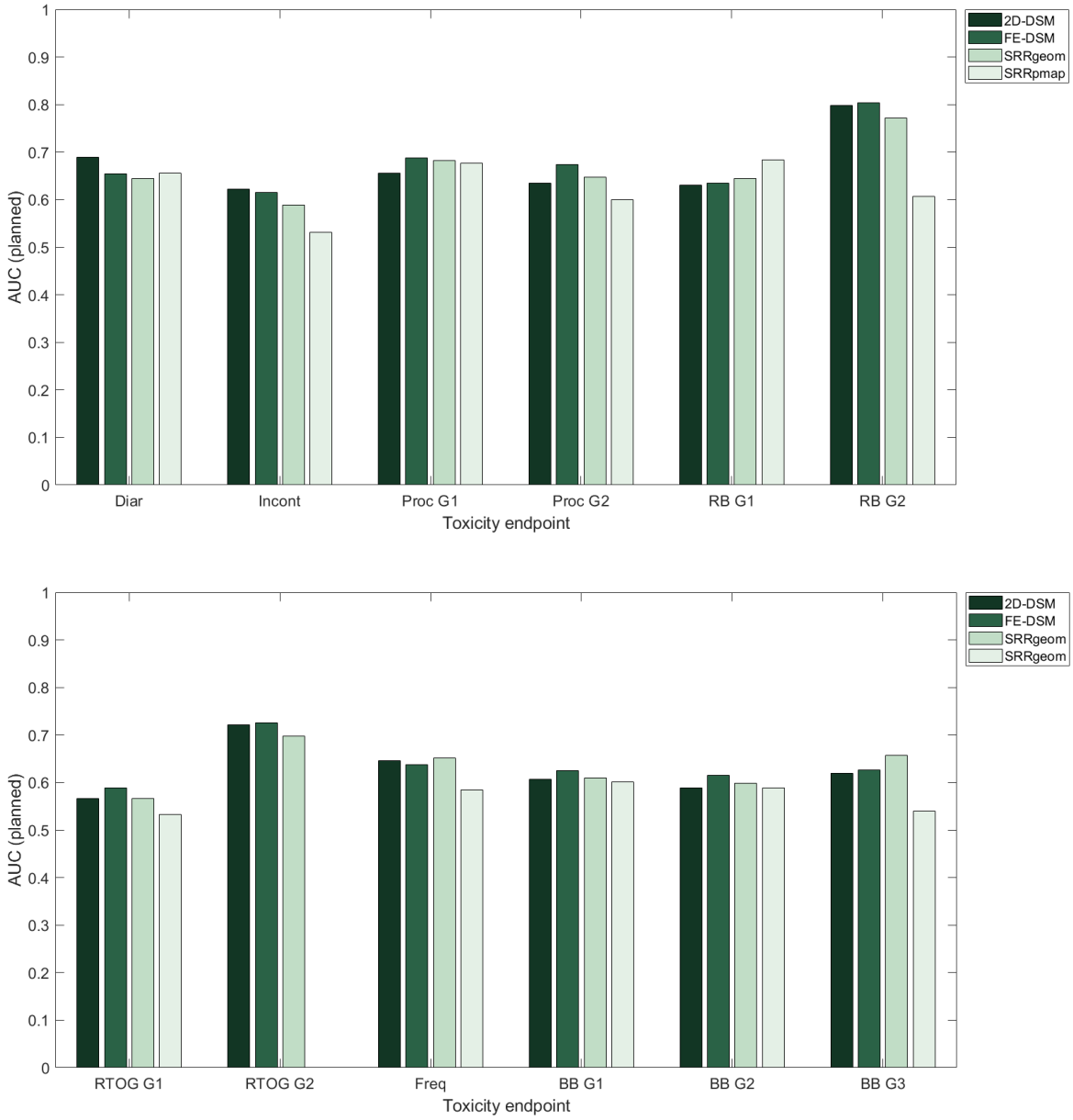


Fig. 6.11 Area under the receiver operator characteristic curve (AUC) for the highest performing dose metric from planned 2D-DSM, FE-DSM, geometric subregion (SRRgeom), and probabilistic subregion (SRRpmap) for 12 toxicity endpoints. Errors have been discussed in corresponding sections.

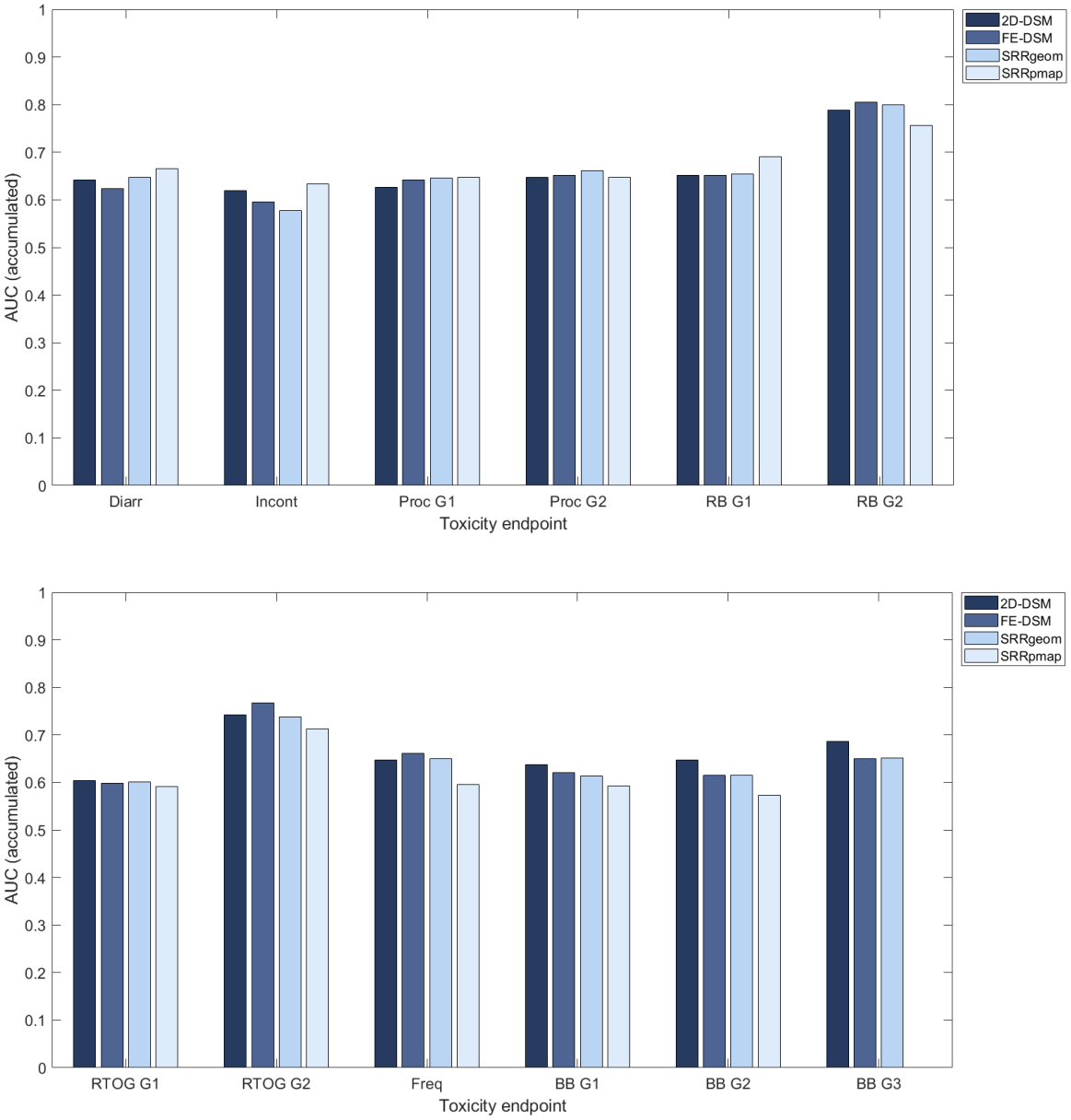


Fig. 6.12 Area under the receiver operator characteristic curve (AUC) for the highest performing dose metric from accumulated 2D-DSM, FE-DSM, geometric subregion (SRRgeom), and probabilistic subregion (SRRpmap) for 12 toxicity endpoints. Errors have been discussed in corresponding sections.

Chapter 7

Predictive models for rectal toxicity

Predictive models are used in radiotherapy to maximise tumour control probability (TCP) and minimise normal tissue complication probability (NTCP). In this chapter, we develop NTCP models specific to our data to address three questions:

1. Is accumulated dose a better predictor of toxicity than planned dose?
2. Does finite element dose accumulation improve the predictive power of NTCP?
3. Can finite element derived subregions improve toxicity prediction?

In previous chapters we explored different methods for parameterisation of planned and delivered dose to the rectal wall in prostate radiotherapy. In Chapter 4, two dimensional dose surface maps (2D-DSM) were generated by uniformly normalising the circumference of the rectal wall, assuming in-plane expansion only. 2D-DSMs were parameterised using equivalent uniform dose (EUD) and DSM dose-widths, the lateral extent of an ellipse fitted to a given isodose level. Chapter 5 describes generating dose surface maps using finite element analysis (FE-DSM), where voxels deform in 3D according to applied boundary conditions and biomechanical properties. For comparative measurement with 2D-DSM results, FE-DSMs were also parameterised using EUD and DSM dose-widths (Chapter 6). Spatial dose information contained within FE-DSMs was further parameterised by identifying rectal subregions at risk using geometric division (SRR_{geom}) and statistical mapping (SRR_{pmap}).

Here, we combine dosimetric parameters from planned and delivered dose with pre-treatment clinical factors to construct multivariate NTCP models for 12 toxicity endpoints. Final model performance indicates whether delivered dose is a better predictor of toxicity than planned dose, whether dosimetric information extracted from FE-DSMs improves model performance with respect to features from 2D-DSMs, and whether subregions reveal areas of heightened radiosensitivity.

The knowledge base of the underpinning statistical methods for constructing NTCP models within this PhD was largely formed by attending the European Society for Radiotherapy & Oncology (ESTRO) school 'Quantitative Methods in Radiation Oncology: Models, Trials and Clinical Outcomes', as well as following approaches described in the literature. The format adopted for documenting model development and validation follows the recommendations of the Transparent Reporting of a multivariate prediction model for Individual Prognosis Or Diagnosis (TRIPOD) Statement [45] where possible.

7.1 Introduction

The aim of radiotherapy treatment planning is to maximise radiation dose to the tumour whilst minimising dose to nearby healthy organs. In prostate radiotherapy, the planning target volume (PTV) is the prostate, and the organs at risk (OARs) include the rectum, bladder, femoral heads. With advances in radiotherapy imaging and hardware, greater conformity and targeting accuracy has enabled dose escalation to the PTV, which has led to improvements in biochemical control of the tumour [84, 85, 130]. However, as a consequence, this can lead to higher dose being received by the healthy OARs. During the iterative radiotherapy treatment planning process, the plan is optimised such that the PTV receives the prescribed radiation dose and the OARs do not exceed certain normal tissue dose thresholds, or constraints. These OAR constraints are generally based on consensus guidelines (e.g. Quantitative Analyses of Normal Tissue Effects in the Clinic, QUANTEC [94]), or the outcomes of randomised clinical trials [10, 54, 148]. The overarching goal is to maximise tumour control probability (TCP) and minimise normal tissue complication probability (NTCP).

NTCP is based on dose-response modelling. OARs generally receive non-homogeneous doses, so the complex 3D dose distribution, often presented as a dose-volume histogram (DVH,) is reduced to a simplified dose metric for NTCP modelling. For example, the proportion of volume receiving a percentage of the prescribed dose or more, the mean dose, and the equivalent uniform dose (EUD), are commonly used. A limitation of these metrics is the loss of spatial dose information due to dimension reduction.

Many studies have investigated the relationship between dose and rectal toxicity using NTCP models[87, 131]. Often, dose has been considered to the entire rectum, but more recent studies have parameterised dose to specific regions of the rectum [34, 35, 61, 115]. This allows potential inhomogeneous intra-organ radiosensitivities to be investigated. Furthermore, various rectal toxicity endpoints have been reported in the literature, and the difference in underlying pathophysiology [70, 126] may be associated with different dose levels. A study by Schaake et al. [140] presented the first multivariate NTCP model considering dose to

specified regions of the rectum, for different toxicity endpoints, as well as clinical prognostic factors. They conclude that different rectal subregions and dose levels are associated with different toxicity effects, which may be useful for plan optimisation to reduce side-effects in future studies.

A common limitation of the published literature is that, to date, rectal NTCP models have been based on planned dose data. This is calculated from the pre-treatment CT scan that represents the anatomy at a static point in time. These data are readily available in radiotherapy as they form the basis of each patient's radiation treatment. However, planned dose does not incorporate the effects of interfraction motion, which introduces deviations to dose parameters [139, 142] and spatial dose patterns [34], and may be more pronounced in hollow structures such as the rectum [140].

Here we present, for the first time, multivariate NTCP models developed for both planned and motion-inclusive accumulated delivered dose to the rectum. The effect of incorporating different spatial dose features of dose surface maps (DSMs) and patient prognostic variables, are investigated for 12 different toxicity endpoints. The approach follows published methodologies [16, 140] for building robust NTCP models in order to demonstrate the merits of different dose parameterisation approaches, and the differences between planned and accumulated dose models.

The risk of toxicity approximately follows a sigmoidal curve when plotted against dose due to radiation dose response relationships. This supports the use of logistic regression to predict the probability of a binary outcome, or the proportion of patients with a given toxicity endpoint. Schakke et al [140] developed multivariable model for rectal bleeding, faecal incontinence and stool frequency based on planned dose volume histograms and found different dose levels and patient characteristics to be associated with each of the discrete endpoints. Following a similar approach, binary logistic regression was selected to develop NTCP models based on VoxTox data. Patient baseline cofactors were included in model development, as well as the spatial dose parameters derived throughout this thesis for planned and accumulated dose.

.

7.2 Material and methods

7.2.1 Patient selection

One hundred and eighty-six prostate cancer patients treated with TomoTherapy between June 2013 and January 2018 were identified from the VoxTox consolidation cohort (as presented

in Section 3.1) based on the availability of a minimum of two year follow-up data available at timepoints 6, 12, and 24 months ($\pm$ 2 months). Twelve toxicity endpoints were considered, as detailed in Table 7.1.

Within each toxicity category, patients were stratified by radiation dose prescription, then according to cumulative toxicity status at 2 years post-radiotherapy, and randomised. Patients were split into training (75%, n = 139), and validation (25%, n = 47) sets such that the ratio of toxicity incidence and prescription proportions were preserved as closely as possible. Using non-random split-sample training and validation sets implies that the final model will be more powerful than internal validation using resampling techniques only [45], and acts to help avoid overfitting of data. External validation was not possible because, as far as we are aware, this is the only dataset to exist with prospectively collected toxicity data, daily image guidance with rectal contours, and daily dose calculation accounting for interfraction motion.

Table 7.1 Voxtox consolidation cohort toxicity incidence rates split into training and validation sets for each of the 12 endpoints under investigation.

	Consolidation cohort (n = 186)		Training set (n = 139)		Validation set (n = 47)	
Toxicity endpoint	n	%	n	%	n	%
Diarrhoea $\geq$ G1	48	25.8	36	25.9	12	25.5
Faecal incontinence $\geq$ G1	32	17.2	24	17.3	8	17.0
Proctitis $\geq$ G1	35	18.8	26	18.7	9	19.1
Proctitis $\geq$ G2	27	14.5	20	14.4	7	14.9
Rectal bleeding $\geq$ G1	62	33.3	46	33.1	16	34.0
Rectal bleeding $\geq$ G2	21	11.3	16	11.5	5	10.6
GI toxicity $\geq$ G1	98	52.7	73	52.5	25	53.2
GI toxicity $\geq$ G2	30	16.1	22	15.8	8	17.0
Stool frequency $\geq$ G1	45	24.2	34	24.5	11	23.4
Bowel bother $\geq$ G1	73	39.2	55	39.6	18	38.3
Bowel bother $\geq$ G2	43	23.1	32	23.0	11	23.4
Bowel bother $\geq$ G3	17	9.1	13	9.4	4	8.5

Patients were prescribed either 74 Gy in 37 fractions (n = 110) or 60 Gy in 20 fractions (n = 76). The decision to combine fractionation regimes was based on the findings of the CHHiP trial [52] where the 74 and 60 Gy trial arms were found to be isoeffective, with no significant difference in rectal toxicity rates. Our data were fit to the CHHiP protocol dose constraints and, assuming equal biologically effective dose, we calculated an α/β ratio of 2.149 for the rectum. Using this α/β, the individual voxels of DSMs for patients prescribed 60 Gy in 20 fractions were corrected to the equivalent dose in 37 fractions (see Section 3.3). By calculating the α/β ratio specific to our study, we can better account for the

Table 7.2 Occurrence of predictor variables in the consolidation cohort, split into training and validation sets. Patient characteristics and clinical history were quoted using the diarrhoea dataset as the representative sample. Baseline incidence rates were quoted from the relevant endpoint where applicable.

Predictor variables	Training (n=139)	Validation (n=47)
Patient characteristics		
Mean age (range), *years*	69 (51-84)	70 (51-82)
Mean Body Mass Index (range), kg/m^2	29 (20-47)	29 (21-39)
Clinical history		
Diabetes	22 (16%)	5 (11%)
Cardiovascular disease	32 (23%)	8 (17%)
Hypertension	65 (47%)	17 (36%)
Inflammatory bowel disease	2 (1%)	2 (4%)
Diverticular disease	15 (11%)	4 (9%)
Haemorrhoids	22 (16%)	6 (13%)
Previous abdominal or pelvic surgery	26 (19%)	13 (28%)
Smoking (range), eq. pack years	14 (0-104)	7 (0-52)
Alcohol (range), av. units/week	12 (0-80)	9 (0-52)
Baseline symptoms		
Diarrhoea	10 (7%)	2 (4%)
Faecal incontinence	1 (1%)	0 (0%)
Proctitis	5 (4%)	1 (2%)
Rectal bleeding	7 (5%)	2 (5%)
GI toxicity	19 (14%)	4 (9%)
Stool frequency	8 (6%)	2 (5%)

inhomogeneous doses received by the rectum, where previous assumptions have been based on a homogeneous dose distributions, as received by the prostate.

One patient's FE model failed from the planning CT data. The reason for this was excessive distortion of a material element, meaning that the deformation of the rectal contour was outwith the capabilities of the model and no solution could be determined. As a consequence, no baseline data would be available to compare planned and accumulated dose, so this patient was removed prior to any analysis. Final numbers are shown in Table 7.2. Four further patients had up to 3 daily FEMs that failed, but rather than excluding these patients, the missing daily delivered dose was substituted with planned dose for the corresponding number of fractions and used to calculate the accumulated delivered dose.

7.2.2 Univariate analysis

7.2.2.1 Pre-treatment factors

Pre-selection of predictor variables was performed prior to developing the model for each toxicity endpoint. Inclusion was based upon previous investigations of rectal toxicity in the literature [87]. Univariate logistic regression analysis was used to test for a relationship between the following patient characteristics and every toxicity endpoint: age, body mass index, alcohol intake (units/week), smoking history (equivalent pack years), and pre-treatment status (1/0) of diabetes, cardiovascular disease, hypertension, inflammatory bowel disease, diverticular disease, haemorrhoids, previous abdominal or pelvic surgery, and presence of baseline toxicity for the endpoint of interest, as shown in Table 7.2. Odds ratios were calculated, and crudely observed effects with $p \leq 0.157$ (Wald test) were selected for inclusion in multivariate modelling [140].

7.2.2.2 Dosimetric parameterisation approaches

Previously in this thesis, planned and accumulated dose to the rectal wall have been parameterised using four discrete approaches:

1. 2D-DSM: EUD and dose-widths (Chapter 4)
2. FE-DSM: EUD and dose-widths (Chapter 6)
3. FE-DSM: EUD of geometrically defined subregions, SRR_{geom} (Chapter 6)
4. FE-DSM: EUD of statistically determined subregions, SRR_{pmap} (Chapter 6)

The results of the receiver operator characteristic (ROC) analysis have been presented in the relevant previous chapters. Dosimetric parameters that demonstrated significant association with toxicity (area under the ROC curve (AUC) ≥ 0.6 and the lower 95% confidence interval ≥ 0.5) were selected for inclusion in multivariate modelling in this chapter. Dose parameters were tested for collinearity using the Pearson coefficient of determination, r^2 and tolerance ($tol = 1 - r^2$), and removed if the variance inflation factor ($VIF = 1/tol$) was > 5, indicating a high degree of collinearity [140]. This affected immediately adjacent dose levels for DSM dose-width measurements. If in the final model, two collinear dose parameters were included in the model, the lesser weighted parameter was removed from the model.

7.2.3 Multivariate analysis

Multivariate logistic regression modelling was performed in SPSS®(IBM®23.0.0.2) for each toxicity endpoint. Separate models were developed for each of the four methods of dose parameterisation described above, for both planned and delivered dose. Covariates were included in the model using stepwise forward variable selection. At each step, the most significant predictor variable was added to the model, and the difference between the model with and without the added predictor variable was assessed using a likelihood ratio test [16]. Covariates were included in the final model where the difference was statistically significant ($p \leq 0.05$). Any covariates resulting in a negative coefficient were retrospectively removed from the analysis, and the model re-generated [140].

Each model was internally validated and corrected using 1000-sample bootstrapping. This is a form of internal validation whereby the model simulation is iteratively run on a randomly resampled proportion of the data. The performance is evaluated at each iteration, effectively reducing the statistical uncertainty in the final model parameters. Model performance was assessed using a Hosmer-Lemeshow test for goodness-of-fit (where a non-significant χ^2 statistic indicates good agreement between expected and observed toxicity rates) and the Nagelkerke R^2 which approximates explained variance.

The output of the binary logistic regression model (Table 7.3) determined the coefficients and covariates defining the parameter S, used in the following NTCP equation [140]:

$$NTCP = \frac{1}{(1+e^{-S})} \tag{7.1}$$

Receiver Operator Characteristic (ROC) analysis was performed to determine the discriminatory ability of each NTCP model with the corresponding toxicity endpoint. The area under the ROC curve (AUC) is equal to 1 for an ideal correlation, and is considered significant when the AUC is ≥ 0.6 with a lower 95% confidence interval (CI) ≥ 0.5 [71].

7.2.4 Validation

The validation set approximately maintains the incidence rates of the different toxicity endpoints as well as the ratios of the different prescription groups of the training set (Table 7.1). However, proportions may be less consistent for patient characteristics, clinical history, and baseline symptoms (Table 7.2). NTCP was calculated for the validation set using the linear regression coefficients obtained from the training set, for each relevant toxicity outcome. Model performance was evaluated using AUCs.

Table 7.3 Output of the binary logistic regression analysis forming the equation for S used in the NTCP equation 7.1. Entries for endpoints or method of dose parameterisation have been excluded where binary logistic regression did not produce a model, or did not produce a model which included a dosimetric variable. Abbreviations: Incont.: faecal incontinence, alc.: average weekly alcohol intake, BL_{tox}: baseline status of tox, dw: dose-width, post.: posterior, quart.: quarter, inf.: inferior, hemi.: hemisphere.

Toxicity	Method	$S_{planned}$	$S_{accumulated}$
Diarrhoea $\geq G1$	2D-DSM	0.030(alc.)+1.871(BL_{diarr})+0.065(35 Gy dw)-7.285	0.031(alc.)+2.109(BL_{diarr})+0.053(35 Gy dw)-6.250
	FE-DSM	0.030(alc.)+1.992(BL_{diarr})+0.063(35 Gy dw)-7.335	0.028(alc.)+2.128(BL_{diarr})+0.068(50 Gy dw)-6.196
	FE-SRR_{geom}	0.027(alc.)+2.068(BL_{diarr})+0.106(post-quart)-5.403	0.028(alc.)+2.213(BL_{diarr})+0.120(post-quart)-5.846
	FE-SRR_{pmap}	0.028(alc.)+1.849(BL_{diarr})+0.051(SRR_{pmap})-4.048	0.028(alc.)+1.933(BL_{diarr})+0.079(SRR_{pmap})-5.564
Incont. $\geq G1$	FE-SRR_{pmap}		0.037(alc.)+0.093(SRR_{pmap})-6.941
Rectal bleeding $\geq G1$	2D-DSM	1.956(BL_{RB})+0.246(EUD)-15.832	1.888(BL_{RB})+0.207(EUD)-13.115
	FE-DSM	1.971(BL_{RB})+0.240(EUD)-15.480	1.927(BL_{RB})+0.073(55 Gy dw)-5.190
	FE-SRR_{geom}	2.243(BL_{RB})+0.053(inf-post-hemi)-2.916	2.293(BL_{RB})+0.059(inf-post-hemi)-3.094
	FE-SRR_{pmap}	2.049(BL_{RB})+0.084(SRR_{pmap})-5.084	1.919(BL_{RB})+0.148(SRR_{pmap})-8.971
Rectal bleeding $\geq G2$	2D-DSM	0.141(60 Gy dw)-9.456	0.124(60 Gy dw)-8.079
	FE-DSM	1.822(BL_{RB})+0.178(60 Gy dw)-12.605	1.917(BL_{RB})+0.167(55 Gy dw)-12.692
	FE-SRR_{geom}	0.250(post-hemi)-14.182	1.175(HTN)+1.934(BL_{RB})+0.310(post-hemi)-17.638
	FE-SRR_{pmap}		1.559(HTN)+0.285(SRR_{pmap})-18.804
GI toxicity $\geq G2$	2D-DSM	1.368(diab.)+0.113(60 Gy dw)-7.759	1.187(HTN)+0.108(45 Gy dw)-9.915
	FE-DSM	1.299(diab.)+1.045(HTN)+0.139(55 Gy dw)-11.305	1.417(diab.)+1.090(HTN)+0.172(50 Gy dw)-14.433
	FE-SRR_{geom}	1.132(HTN)+0.173(post-quart)-8.600	1.324(diab.)+0.224(post-quart)-10.079
	FE-SRR_{pmap}		0.125(SRR_{pmap})-7.935
Stool frequency $\geq G1$	2D-DSM	2.460(BL_{freq})+0.050(50 Gy dw)-4.541	2.575(BL_{freq})+0.065(60 Gy dw)-4.436
	FE-DSM		2.667(BL_{freq})+0.073(60 Gy dw)-5.291
	FE-SRR_{geom}	2.501(BL_{freq})+0.103(post-quart)-5.079	2.565(BL_{freq})+0.105(post-quart)-5.056
Bowel bother $\geq G1$	2D-DSM	0.040(40 Gy dw)-3.577	0.069(60 Gy dw)-3.658
	FE-DSM	0.045(40 Gy dw)-4.155	0.049(55 Gy dw)-3.428
	FE-SRR_{geom}		0.119(right-quart)-7.229

7.3 Results

7.3.1 Logistic regression output

The output of the multivariate binary logistic regression modelling is presented in Table 7.3. Models that did not produce an output based on the forward selection criteria were omitted, as were models that did not include a significant dose metric contribution. This may occur when patient characteristics or clinical cofactors were most strongly associated with the toxicity outcome, and the addition of a dose contribution into the model was of no additional benefit. In total, 7 toxicity endpoints resulted in statistically significant models, based on different dose parameterisation approaches. Final models were internally validated using 1000-sample bootstrapping. Performance indicators of models with the greatest discriminating ability for planned and accumulated dose metrics are shown in Table 7.4. NTCPs were calculated using Equation 7.1 by inserting patient data into equations for S. An example NTCP curve for rectal bleeding is shown in Figure 7.1. AUCs were calculated for all models, and the top performing models for each endpoint are shown in Figure 7.2a, tested on training set data, and Figure 7.2b, tested on validation set data.

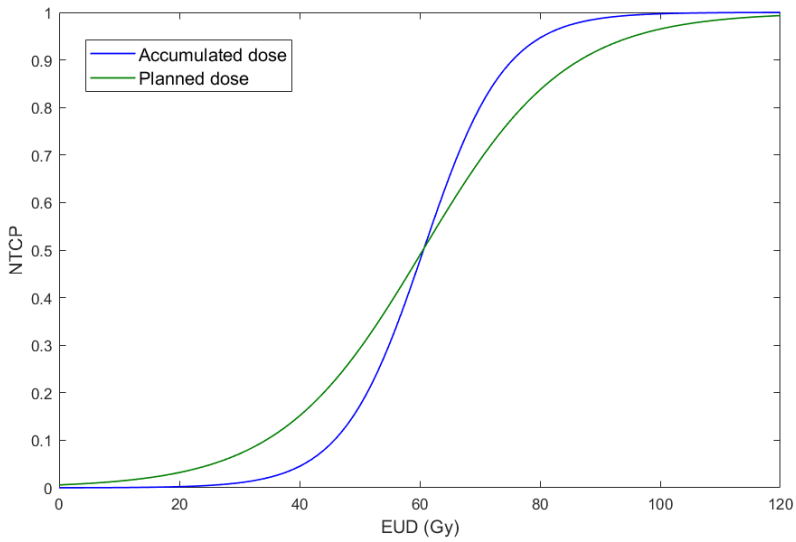


Fig. 7.1 Example normal tissue complication probability (NTCP) model calculated for CTCAE rectal bleeding $\geq$ G2, based upon planned (green) and accumulated (blue) EUD of the probabilistic rectal subregions at risk ($SRR_{0.05}$) from the finite element based dose surface map (FE-DSM). The resulting NTCP curve is steeper when using accumulated dose, meaning that a small change in dose has a greater effect on NTCP.

7.3.2 Planned versus accumulated dose

Binary logistic regression produced 23 models using accumulated dose data, compared to 18 from planned dose data, across 7 toxicity endpoints using up to four different methods of dose parameterisation (Table 7.3). Of the 23 models developed using training set data, the AUCs of NTCP models generated from accumulated dose metrics were greater than those from planned dose metrics for 15/23 models (65 %). The dosimetric parameters included in the planned dose model were generally the same as or close to those selected for the equivalent accumulated dose model. Of particular note is the repeated occurrence of subregions of the posterior rectal wall (lower dose regions) selected for models based upon geometric subregions, across different endpoints. Testing the NTCP models using the validation set produced greater AUCs when using accumulated dose metrics for 16/23 models (70 %). The model resulting in the greatest AUC for each endpoint is shown in Table 7.4, alongside measures of performance of logistic regression output. Per endpoint, the Nagelkerke R^2 was greater by an average of 2.5% (range -1.4, 6.9) when comparing the best performing accumulated dose based model with planned, indicating an overall increase in explained variance when using accumulated dose models. The logistic regression results for all endpoints shown in Table 7.4 found the Hosmer-Lemeshow test to be non-significant, indicating good agreement between expected and observed toxicity rates for planned and

Table 7.4 Performance of multivariate binary logistic regression models for 7 toxicity endpoints. Models were internally validated using a 1000 sample bootstrapping procedure. Plan./acc. indicates whether model was developed using planned or accumulated dose-surface maps (DSMs). Method and metric indicate DSM calculation and parameterisation approach. Performance indicators Nagelkerke R^2 and Hosmer-Lemeshow (H-L) χ^2 and p-value describe explained variance and goodness-of-fit, respectively, where a non-significant χ^2 statistic indicates good agreement between expected and observed toxicity rates.

Toxicity endpoint	plan./acc.	Method	Metric	Nagelkerke R^2	H-L χ^2	H-L p-value
Diarrhoea	plan.	2D-DSM	35 Gy dw	0.220	4.094	0.849
$\geq G1$	acc.	2D-DSM	35 Gy dw	0.206	7.261	0.509
Incont.$\geq G1$	acc.	FE-SRR$_{pmap}$	EUD(SRR$_{0.05}$)	0.130	11.576	0.171
Rectal bleeding	plan.	FE-SRR$_{pmap}$	EUD(SRR$_{0.05}$)	0.172	9.496	0.302
$\geq G1$	acc.	FE-SRR$_{pmap}$	EUD(SRR$_{0.05}$)	0.190	5.054	0.752
Rectal bleeding	plan.	FE-DSM	60 Gy dw	0.258	1.765	0.987
$\geq G2$	acc.	FE-SRR$_{geom}$	EUD(post-hemi)	0.290	7.378	0.496
GI toxicity	plan.	FE-DSM	55 Gy dw	0.228	14.012	0.081
$\geq G2$	acc.	FE-DSM	50 Gy dw	0.297	14.613	0.067
stool frequency	plan.	FE-SRR$_{geom}$	EUD(post-quart)	0.159	6.157	0.63
$\geq G1$	acc.	FE-DSM	60 Gy dw	0.172	5.242	0.731
Bowel bother	plan.	FE-DSM	40 Gy dw	0.054	10.117	0.257
$\geq G1$	acc.	2D-DSM	60 Gy dw	0.091	4.833	0.775

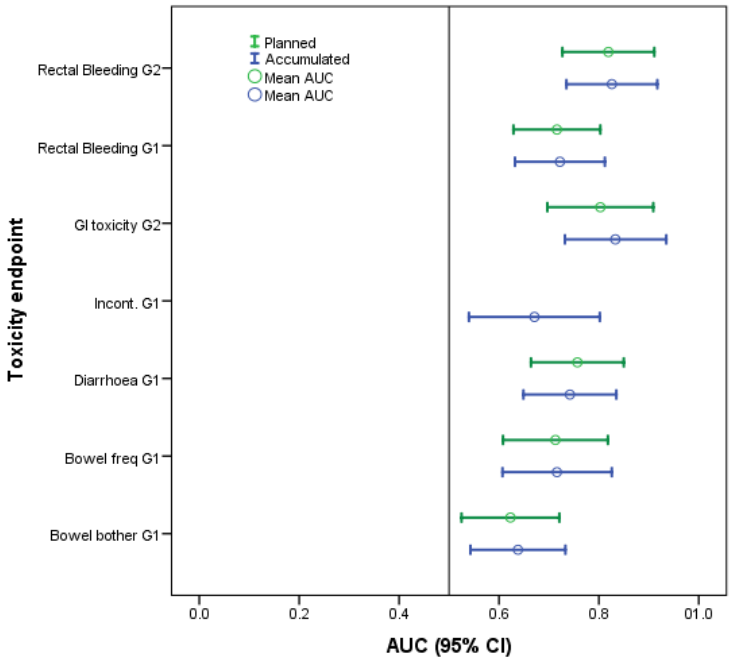
accumulated dose based models. A student t-test found differences in NTCP distributions to be significant ($p \leq 0.01$) between patient groups with and without toxicity.

7.3.3 2D-DSM versus FE-DSM

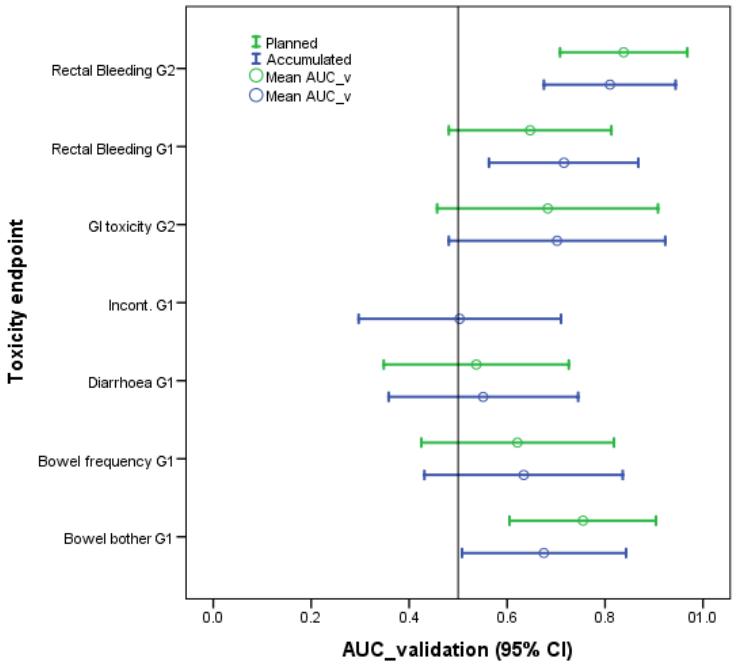
When considering different methods of calculating dose to the rectal wall, NTCP models based on FE-DSM dose metrics were found to have higher discriminative ability than those based on 2D-DSM dose metrics for 5/7 toxicity endpoints (71 %) in the training set, and 6/7 NTCP models (86 %) in the validation set.

7.3.4 Dose metric comparison

Considering the 5/7 toxicity endpoints where FE-DSM-based NTCP models produced greater AUCs than 2D-DSM-based NTCP models in the training set, of these, the greatest AUCs were found for faecal incontinence $\geq G1$ and rectal bleeding $\geq G1$ using FE-SRR$_{pmap}$ (2/5), for rectal bleeding $\geq G2$ using FE-SRR$_{geom}$ (1/5), and for GI toxicity $\geq G2$ and stool frequency $\geq G1$ using FE-DSM dose-widths (2/5). In the validation set, rectal bleeding $\geq G1$ still produced the greatest AUC using FE-SRR$_{pmap}$ (1/5), but all others produced the greatest AUC from FE-DSM dose-widths (4/5).



(a) Training set data (n = 139), used in model generation.



(b) Validation set data (n = 47), not used in generating models.

Fig. 7.2 Discriminative ability of NTCP model and respective toxicity outcome presented using area under the receiver operator characteristic curve (AUC) and 95% confidence intervals (CI). Planned and accumulated models plotted correspond to those with the highest discriminative ability per endpoint, as listed in Table 7.4, and detailed in Table 7.3. Note no NTCP model was generated for incontinence from planned dose.

7.4 Discussion

For the first time, NTCP models have been developed for different rectal toxicity endpoints based on pre-treatment variables and dosimetric parameters from accumulated delivered dose. Robust methodologies have been followed to compare the predictive power of NTCP models optimised for accumulated dose, with those optimised for planned dose. Overall, results suggest that models based on accumulated dose have greater predictive power than those based on planned dose metrics.

Different approaches to dose parameterisation have been investigated throughout this thesis and compared using NTCP modelling in this chapter. Generally, accumulating dose using FE-DSMs was found to be more predictive of toxicity. FE modelling allows dose to be tracked per voxel according to biomechanical properties which is more anatomically representative [21] than the more simple in-plane uniform normalisation approach used for 2D-DSMs. However, in terms of parameterising dose from the FE-DSMs, there was no consensus agreement as to the best method selected by the logistic regression model across all endpoints. This may be due to inherent differences in underlying toxicity pathophysiologies. With increasing sophistication in identifying SRRs, there was no clear generalised gain in terms of discriminative ability, but some benefit was observed on an endpoint-by-endpoint basis. Further investigation into methods of extracting meaningful dose information from FE-DSMs may be beneficial. However, this result supports the findings in the literature that the underlying pathophysiology is endpoint-specific [12, 126], and as such, dose to specific regions of the rectum may be more strongly associated with different toxicities, hence the variations in dose parameters being found most significant in NTCP modelling.

Strong associations were observed for rectal bleeding $\geq G2$ with NTCP models developed using both accumulated (AUC 0.826, 95% CI [0.735, 0.917]) and planned (AUC 0.819, 95% CI [0.727, 0.911]) dose metrics. The accumulated NTCP model found the dose to the posterior-hemi subregion of the rectal wall to be most significant, which put more weighting on lower doses (SRR EUD median 46 Gy, range 27-55 Gy, compared with EUD median 61 Gy, range 45-65 Gy, of entire rectal wall). The planned dose metric found to be the most significant was the FE-DSM 60 Gy dose-width. Although there are no results in the literature using FE-DSMs for planned dose data, findings based upon 2D-DSMs of planned dose to the rectal wall have found dose-widths close to 60 Gy to be predictive of rectal bleeding [27, 70, 142, 115, 34, 49]. This was supported by the outcome of the logistic regression for the planned 2D-DSM (and accumulated 2D-DSM, but the final NTCP model was found to be more discriminative of rectal bleeding when using FE-DSM 60 Gy dose-width). Both planned and accumulated NTCP models were found to be highly associated with rectal bleeding during model validation $\geq G2$ (AUC 0.838 [95% CI: 0.708, 0.968]

and AUC 0.810, [95% CI: 0.675, 0.944], respectively). Rectal bleeding is one of the more commonly reported endpoints in the literature as symptoms are less subjective that other endpoints. In particular, grade 2 rectal bleeding by definition requires medical intervention, so this result could be useful in clinical decision making in radiotherapy treatment planning as well as adaptive radiotherapy.

Rectal bleeding $\geq G1$ was most strongly associated with rectal subregions identified through statistical analysis (Chapter 6). The SRR_{pmap} regions were defined where differences in pixel dose were most significant between patients with and without toxicity. For rectal bleeding $\geq G1$ this roughly incorporated a lateral band around the circumference of the lower third of the rectal wall, and a smaller cluster around the mid-right of the rectum. Validation results were significant for accumulated SRR_{pmap} based NTCP model, but 95% CI extended below 0.5 for the equivalent planned NTCP model. Findings suggest that different regions of the rectal wall associated with $\geq G1$ and $\geq G2$.

The NTCP model developed for GI toxicity $\geq G2$ based on the accumulated FE-DSM 50 Gy dose-width produced the greatest AUC across all endpoints (AUC 0.833 [95% CI: 0.732, 0.935]). From the planned dose based NTCP model, the FE-DSM 55 Gy dose-width was found to be the most significant dose metric (AUC 0.803 [95% CI: 0.697, 0.909]). However, the 95% CIs calculate for AUCs during model validation were wide and extended below the 0.5 threshold (accumulated AUC 0.702, [95% CI: 0.481, 0.923], planned AUC 0.683 [95% CI: 0.457, 0.908]).

All remaining NTCP models were found to have associations with the relevant toxicity endpoint. Considering the significant dose metric included in NTCP modelling following logistic regression, results from both accumulated and planned dose suggest that there are features of low, intermediate, and high doses associated with different toxicity endpoints. This has been an effect reported in the literature based on planned dose, but it is supported here by accumulated dose results.

The multivariate model was generated using forward variable selection based on likelihood ratios, which carries an associated risk of overfitting [16]. However, to address this, the number of variables was kept low (between 1 and 3), with only one dosimetric variable appearing in each equation for S shown in Table 7.3. Furthermore, because the data were considered as TRIPOD Type 2b non-random split-sample development and validation, not only was the model internally validated using bootstrapping, but the predictive power of the final model was further interrogated using data not considered during model development. This approach gives strength to the model design when compared with internal validation only.

Testing the NTCP models using the validation set resulted in wider confidence intervals across all models. This is expected due to the reduced number of samples in the validation set, and may also be due to differences in rates of baseline variables. The models found to have significant discriminative ability following validation were: rectal bleeding $\geq G2$ and bowel bother $\geq G1$, for both planned and accumulated dose metrics, and rectal bleeding $\geq G1$ for accumulated dose only. Notably, GI toxicity and stool frequency produced AUCs greater than 0.6, but the widths of the 95 % CIs extended below 0.5. However, the same cutoff thresholds are not applicable because the performance of the validation set will generally not be as good as the training set. Accumulated dose produced AUCs greater than planned for 5 out of 7 endpoints.

Although the intuitive conclusion is that accumulated dose is a stronger predictor of toxicity than planned dose because it incorporates the effects of interfraction motion, and that FE-DSMs are a better tool for dose accumulation than 2D-DSMs, the differences in AUC when translated into NTCP models are small. Across a population, this suggests a small but measurable effect. This can be considered a positive outcome when put in the context of the low event rates being investigated, and the inherent noise in toxicity reporting. Meaningfully quantifying accumulated delivered dose is not yet fully understood, and further research is recommended because the DVH approaches used for describing planned dose are not appropriate for dose accumulation. However, the results presented here suggest that for most endpoints, there is a small but non significant improvement in calculating NTCP using accumulated dose rather than planned. Although we have detected large deviations in daily rectal contours from the planning CT baseline (Chapter 5), the most extreme effects are blurred when considering the entire course. Work is ongoing to improve the accuracy of dose accumulation techniques, including autosegmentation and FE refinement. Results for NTCP modelling suggest that planned dose is an adequate surrogate for delivered dose over the course of a radiotherapy treatment.

Where the small but measurable differences reported here may begin to influence a patient's treatment are for ultra-hypofractionated treatments on standard linacs or cyberknife, where more dose is being delivered per fraction and α/β ratios come into account in terms of biological dose-effects. Hypofractionated treatments are anticipated to be delivered on the MR-linac (magnetic-resonance linear accelerator). Another consideration is proton beam radiotherapy whereby small effects of intra/interfraction motion can have a large impact on the patient's dosimetry due to the inherent high dose gradients and range of narrow beams. If patients were treated with hypofractionated proton treatment this would also be a consideration.

NTCP models have been generated to compare predictive power between planned and accumulated dose metrics, with several clinical endpoints. Models developed on a training set and internally validated were found to be strongly associated with each endpoint investigated. Each model was tested using a validation set of patients who had not been considered during model development. Five of thirteen models were successfully validated, but other validation results may have suffered from insufficient numbers in the validation set. Generally, NTCP models based on accumulated dose were more strongly associated with toxicity than those based on planned dose, but the benefits were marginal. Planned dose may be adequately representative of accumulated dose for the course of radiotherapy, and the models developed here may contribute new information on subregions associated with different toxicity endpoints. In particular, some endpoints were found to be associated with low to intermediate dose levels, not only doses approaching the prescription value, and spatial distribution of dose were found to be significant across all NTCP models. By furthering our knowledge of geometrical dose distributions and the corresponding associations with toxicity outcomes, we can improve our understanding of the pathogenesis of toxicity, and therefore inform development of treatments. We envisage that the NTCP models developed in this work will be incorporated into decision-making tools in adaptive radiotherapy, with the aim of reducing the incidence of radiation-induced toxicity and improving precision tumour targeting.

Chapter 8

Conclusions and future work

8.1 Conclusions

The hypothesis of the VoxTox research programme is that delivered dose is a better predictor of toxicity than planned dose. The multidisciplinary project has brought together technical contributions from bioengineering, medical physics, oncology, applied mathematics, and high energy physics, to enable the calculation of delivered dose from daily image guidance radiotherapy scans. The research conducted within this thesis focuses specifically on the calculation of motion-inclusive delivered dose to the rectum in prostate radiotherapy. Through collaborative working, images being routinely acquired for the purposes of patient positioning, can now be exploited to explore delivered dose-toxicity correlations. Within this PhD research: different approaches to interrogating delivered dose were investigated; a biomechanical finite element (FE) rectal simulation was introduced; voxel-level dose analysis was performed; and multivariate normal tissue complication probability (NTCP) models were developed, optimised, and compared between planned and accumulated dose.

8.1.1 Quantifying spatial aspects of delivered dose to the rectum

Two-dimensional dose-surface maps (DSMs) were used to represent dose to the rectal wall so that spatial features of the dose distribution could be preserved during accumulation of daily delivered dose. Equivalent uniform dose (EUD) and spatial dose-width metrics were used to parameterise both planned and accumulated dose.

Dose-toxicity analysis was performed on two separate patient cohorts. Results from both cohorts found accumulated dose to have stronger associations with several toxicity endpoints: grade 2 proctitis, grade 1 rectal bleeding, grades 1 & 2 GI toxicity, stool frequency, and all

3 investigated levels of bowel bother. This was the first reporting of delivered dose being linked with toxicity using image guidance scans from a large patient cohort.

8.1.2 Biomechanical finite element modelling

Biomechanical finite element modelling (FEM) was introduced into the VoxTox workflow for accumulating dose to the rectum. This addressed several limitations associated with the original DSM approach in terms of degrees of freedom for rectal motion. The model was optimised for patient data, and provided a more anatomically-representative solution, allowing accurate accumulation of dose at the voxel level.

8.1.3 FEM motion-dose sensitivity study

FEM was used to simulate different magnitudes and scenarios of rectal motion based on measured patient data. The outcome was to quantify the sensitivity of dose with rectal motion, and interpret this in terms of NTCP implications. Improving the understanding of how dose varies with motion may influence clinical considerations and imaging protocols in the future.

8.1.4 FEM for voxelwise analysis of dose

FE-calculated DSMs were generated for further dose-toxicity investigations in a large patient cohort. Overall, the introduction of FE-based dose calculation led to improved toxicity correlations for both planned and accumulated dose, with accumulated dose tending to perform better. Beyond EUD and dose-width metrics, geometric and probabilistic approaches to defining rectal subregions were investigated. However, increasing resolution of dose parameterisation did not always correspond to improvements in toxicity prediction, despite seeming to reveal intra-organ areas of radiosensitivity. Optimal spatial dose metrics varied across different toxicity endpoints, suggesting differences in underlying pathophysiologies.

8.1.5 NTCP modelling

Multivariate NTCP models were developed based on spatial dose parameters from planned and accumulated DSMs, combined with clinical risk factors. Seven models were considered significantly discriminative of toxicity; grade 1 & 2 rectal bleeding, grade 2 GI toxicity, faecal incontinence, diarrhoea, stool frequency, and grade 1 bowel bother. Models were validated on a separate test cohort. For 6 of 7 endpoints, models based on accumulated dose performed better. However, differences with planned dose-based models were small, and

further research is recommended into extracting clinically useful dosimetric information from delivered dose.

The small detected differences between accumulated and planned dose may be due to limitations within VoxTox calculation systems, or perhaps, in the current era of image guided radiotherapy, planned dose is an adequate surrogate for delivered dose and is in fact a good predictor of radiobiological response.

8.1.6 Research outcomes

The outcomes and models developed within this thesis will help to advance and improve the understanding of the links between dose to the rectum and toxicity in prostate radiotherapy. For the first time, it has been demonstrated that delivered dose can be a better predictor of rectal toxicity than planned dose. Different approaches to dose accumulation and parameterisation metrics have been presented and translated to NTCP models for different toxicity endpoints. Rectal motion has been quantified and modelled within a sensitivity study to demonstrate the relationship between motion, dose, and NTCP. A summary of the research outcomes from this thesis are included in Appendix B. Ultimately, it is anticipated that this research will contribute towards advancing systems for adaptive radiotherapy with the aim of reducing the incidence of rectal toxicity for prostate cancer patients.

8.2 Future work and applications

8.2.1 Machine learning

The approaches applied for parameterising DSMs in this thesis have been adapted from methods presented in the literature for planned dose data [27, 64]. Whilst these metrics have indicated an advantage to using accumulated dose for dose-toxicity investigation over planned dose, the differences are not as distinct as expected, and further research into parameterisation approaches specific to accumulated dose is recommended. Machine learning has already been shown to have some application in this area [177] and may prove to be a useful tool for improving autosegmentation or voxel-level dose-toxicity analysis.

8.2.2 Voxel-level dose analysis

Voxelwise analysis of the anatomy is becoming an area of increased interest [61, 59, 105]. At the recent congress of the European Society for Radiotherapy & Oncology (ESTRO), a working group was assembled of representatives from research groups sharing common

interests. I was fortunate to be invited to sit on the working group and look forward to working collectively to put together guidelines and recommendations for voxelwise analysis to be published in the future.

8.2.3 Rectal spacers

With increasing evidence demonstrating associations between spatial dose patterns and rectal toxicity, research outcomes could complement the emerging use of rectal spacers in prostate radiotherapy [62, 147, 160]. Physical spacers are surgically implanted to insert a gel-based or solidifying barrier between the prostate and rectum, subsequently shifting the rectum away from the high dose region. Evidence has shown a reduction in rectal toxicity of 8-12% [160] and the procedure is being trialled by the NHS. However, additional considerations include any contraindications to general anaesthetic, surgical facilities and availability, and the additional cost, reported as over 2000 USD per patient [80]. The research outcomes from this PhD thesis could be used to help identify patients who would benefit most from rectal spacers, and those patients unlikely to benefit from the procedure due to having low risk of developing toxicity.

8.2.4 Latest technologies

Accumulated delivered dose is becoming an increasingly important consideration for proton beam therapy [2]. Accuracy and precision are of utmost importance because higher doses of radiation are deposited to a focal spot due to the nature of the Bragg peak. Small changes in anatomy can lead to greater differences between planned and delivered dose, and subsequently NTCP, than previously observed for photon based radiotherapy. Accumulated delivered dose has also become of interest to the adaptive workflow of the MR-linac [144]. Treatment has commenced on the first two units installed in the UK [157], and the market is increasing and advancing. The MR-linac was designed for the purpose of image-based adaptive workflow, so implementation of online systems for toxicity monitoring would be a logical addition.

The research outcomes of this PhD thesis have potential implications for directly improving the adaptive workflow for TomoTherapy patients. The systems may also have useful applications for conventional linear accelerators, Cyberknife, MR-linac and proton beam therapy, as well as in delivery techniques such as ultra-hypofractionated radiotherapy. The ultimate aim is that these tools will be used in the development of clinical decision-making tools for adaptive radiotherapy that will benefit the patient by minimising the risk of radiation-induced rectal toxicity.

References

[1] (1995). LENT SOMA tables table of contents. *Radiotherapy and Oncology*, 35(1):17–60.

[2] Abe, Y., Kadoya, N., Arai, K., Takayama, Y., Kato, T., Kimura, K., Ono, T., Nakamura, T., Wada, H., Kikuchi, Y., and Jingu, K. (2017). Effect of DIR uncertainty on prostate passive-scattering proton therapy dose accumulation. *Physica Medica*, 39:113–120.

[3] Acosta, O., Drean, G., Ospina, J. D., Simon, A., Haigron, P., Lafond, C., and de Crevoisier, R. (2013). Voxel-based population analysis for correlating local dose and rectal toxicity in prostate cancer radiotherapy. *Physics in Medicine and Biology*, 58(8):2581.

[4] Adamus-Górka, M., Mavroidis, P., Lind, B. K., and Brahme, A. (2011). Comparison of dose response models for predicting normal tissue complications from cancer radiotherapy: Application in rat spinal cord. *Cancers*, 3:2421–2443.

[5] Al-Mayah, A., Moseley, J., Velec, M., and Brock, K. (2008a). *Effect of Friction and Material Compressibility on Deformable Modeling of Human Lung*, pages 98–106. Springer Berlin Heidelberg, Berlin, Heidelberg.

[6] Al-Mayah, A., Moseley, J., Velec, M., and Brock, K. (2011). Toward efficient biomechanical-based deformable image registration of lungs for image-guided radiotherapy. *Physics in Medicine and Biology*, 56(15):4701.

[7] Al-Mayah, A., Moseley, J. L., and Brock, K. K. (2008b). Contact surface and material nonlinearity modeling of human lungs. *Physics in medicine and biology*, 53(1):305–17.

[8] Al Mayah, A. E. (2018). *Biomechanics of Soft Tissues: Principles and Applications*. CRC Press.

[9] Alterovitz, R., Goldberg, K., Pouliot, J., Hsu, I.-C. J., Kim, Y., Noworolski, S. M., and Kurhanewicz, J. (2006). Registration of MR prostate images with biomechanical modeling and nonlinear parameter estimation. *Medical Physics*, 33(2):446–454.

[10] Aluwini, S., Schimmel, E., Pos, F., van der Toorn, P. P., Dirkx, M., de Jager, H., van Lin, E., Alemayehu, W. G., Incrocci, L., Krol, S., and Heijmen, B. (2015). Hypofractionated versus conventionally fractionated radiotherapy for patients with prostate cancer (HYPRO): acute toxicity results from a randomised non-inferiority phase 3 trial. *The Lancet Oncology*, 16(3):274–283.

[11] Anderson, N. S., Yu, J. B., Peschel, R. E., and Decker, R. H. (2011). A significant decrease in rectal volume and diameter during prostate IMRT. *Radiotherapy and Oncology*, 98(2):187–191.

[12] Aprile, G., Rihawi, K., De Carlo, E., and Sonis, S. T. (2015). Treatment-related gastrointestinal toxicities and advanced colorectal or pancreatic cancer: A critical update. *World Journal of Gastroenterology*, 21(41):11793–11803.

[13] Arcangeli, S., Strigari, L., Gomellini, S., Saracino, B., Petrongari, M. G., Pinnarò, P., Pinzi, V., and Arcangeli, G. (2012). Updated results and patterns of failure in a randomized hypofractionation trial for high-risk prostate cancer. *International Journal of Radiation Oncology Biology Physics*, 84(5):1172–1178.

[14] Barnett, G. C., West, C. M. L., Dunning, A. M., Elliott, R. M., Coles, C. E., Pharoah, P. D. P., and Burnet, N. G. (2009). Normal tissue reactions to radiotherapy: towards tailoring treatment dose by genotype. *Nature Reviews Cancer*, 9(2):134–142.

[15] Bates, A. M., Scaife, J. E., J Tudor, G. S., Jena, R., Romanchikova, M., Dean, J. C., Phys, M., F Hoole, A. C., D Simmons, M. P., Burnet, N., and Madeleine Bates, A. (2013). Image guidance protocols: balancing imaging parameters against scan time. *Br J Radiol*, 86.

[16] Beasley, W., Thor, M., McWilliam, A., Green, A., Mackay, R., Slevin, N., Olsson, C., Pettersson, N., Finizia, C., Estilo, C., Riaz, N., Lee, N. Y., Deasy, J. O., and van Herk, M. (2018). Image-based Data Mining to Probe Dosimetric Correlates of Radiation-induced Trismus. *International Journal of Radiation Oncology*Biology*Physics*, 102(4):1330–1338.

[17] Benjamini, Y. and Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)*.

[18] Boubaker, M. B., Haboussi, M., Ganghoffer, J.-F., and Aletti, P. (2009). Finite element simulation of interactions between pelvic organs: Predictive model of the prostate motion in the context of radiotherapy. *Journal of Biomechanics*, 42(12):1862 – 1868.

[19] Boubaker, M. B., Haboussi, M., Ganghoffer, J.-F., and Aletti, P. (2015). Predictive model of the prostate motion in the context of radiotherapy: A biomechanical approach relying on urodynamic data and mechanical testing. *Journal of the Mechanical Behavior of Biomedical Materials*, 49:30 – 42.

[20] Brock, K., Sharpe, M., Dawson, L., Kim, S., and Jaffray, D. (2005). Accuracy of finite element model-based multi-organ deformable image registration. *MEDICAL PHYSICS*, 32(6):1647–1659.

[21] Brock, K. E. (2014). *Image Processing in Radiation Therapy*. CRC Press.

[22] Brock, K. K., Dawson, L. A., Sharpe, M. B., Moseley, D. J., and Jaffray, D. A. (2006). Feasibility of a novel deformable image registration technique to facilitate classification, targeting, and monitoring of tumor and normal tissue. *International Journal of Radiation Oncology*Biology*Physics*, 64(4):1245 – 1254.

[23] Brock, K. K., Hawkins, M., Eccles, C., Moseley, J. L., Moseley, D. J., Jaffray, D. A., and Dawson, L. A. (2008a). Improving image-guided target localization through deformable registration. *Acta Oncologica*, 47(7):1279–1285.

[24] Brock, K. K., Nichol, A. M., Ménard, C., Moseley, J. L., Warde, P. R., Catton, C. N., and Jaffray, D. A. (2008b). Accuracy and sensitivity of finite element model-based deformable registration of the prostate. *Medical Physics*, 35(9):4019–4025.

[25] Buettner, F., Gulliford, S. L., Webb, S., and Partridge, M. (2009a). Using dose-surface maps to predict radiation-induced rectal bleeding: a neural network approach. *Physics in Medicine and Biology*, 54(17):5139.

[26] Buettner, F., Gulliford, S. L., Webb, S., and Partridge, M. (2011). Modeling late rectal toxicities based on a parameterized representation of the 3d dose distribution. *Physics in Medicine and Biology*, 56(7):2103.

[27] Buettner, F., Gulliford, S. L., Webb, S., Sydes, M. R., Dearnaley, D. P., and Partridge, M. (2009b). Assessing correlations between the spatial distribution of the dose to the rectal wall and late rectal toxicity after prostate radiotherapy: an analysis of data from the MRC RT01 trial (ISRCTN 47772397). *Physics in Medicine and Biology*, 54(21):6535.

[28] Burman, C., Kutcher, G., Emami, B., and Goitein, M. (1991). Fitting of normal tissue tolerance data to an analytic function. *International Journal of Radiation Oncology*Biology*Physics*, 21(1):123 – 135.

[29] Burnet, N., Adams, E., Fairfoul, J., Tudor, G., Hoole, A., Routsis, D., Dean, J., Kirby, R., Cowen, M., Russell, S., Rimmer, Y., and Thomas, S. J. (2010). Practical Aspects of Implementation of Helical Tomotherapy for Intensity-modulated and Image-guided Radiotherapy. *Clinical Oncology*, 22(4):294–312.

[30] Burnet, N. G., Scaife, J., Romanchikova, M., Thomas, S., Bates, A., Wong, E., Noble, D., Shelley, L., Bond, S., Forman, J., Hoole, A., Barnett, G., Brochu, F., Simmons, M., Jena, R., Harrison, K., Yeap, P. L., Drew, A., Silvester, E., Elwood, P., Pullen, H., Sultana, A., Seah, S., Wilson, M., Russell, S., Benson, R., Rimmer, Y., Jefferies, S., Taku, N., Gurnell, M., Powlson, A., Schönlieb, C.-B., Cai, X., Sutcliffe, M., and Parker, M. (2017). Applying physical science techniques and CERN technology to an unsolved problem in radiation treatment for cancer: the multidisciplinary 'VoxTox' research programme. *CERN IdeaSquare Journal of Experimental Innovation*, 1(1):3.

[31] Buyyounouski, M. K., Price, R. A., Harris, E. E. R., Miller, R., Tomé, W., Schefter, T., Parsai, E. I., Konski, A. A., and Wallner, P. E. (2010). Stereotactic Body Radiotherapy for Primary Management of Early-Stage, Low- to Intermediate-Risk Prostate Cancer: Report of the American Society for Therapeutic Radiology and Oncology Emerging Technology Committee. *Int. J. Radiation Oncology Biol. Phys.*, 76(5):1297–1304.

[32] Cancer Research UK (2018). Prostate cancer statistics. http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/prostate-cancer. Accessed: May 2018.

[33] Carter, T. J., Sermesant, M., Cash, D. M., Barratt, D. C., Tanner, C., and Hawkes, D. J. (2005). Application of soft tissue modelling to image-guided surgery.

[34] Casares-Magaz, O., Muren, L. P., Moiseenko, V., Petersen, S. E., Pettersson, N. J., Høyer, M., Deasy, J. O., and Thor, M. (2017). Spatial rectal dose/volume metrics predict patient-reported gastro-intestinal symptoms after radiotherapy for prostate cancer. *Acta Oncologica*, 56(11):1507–1513.

[35] Casares-Magaz, O., Muren, L. P., Pettersson, N., Thor, M., Hopper, A., Knopp, R., Deasy, J. O., Vaeth, M., Einck, J., and Moiseenko, V. (2018). A case-control study using motion-inclusive spatial dose-volume metrics to account for genito-urinary toxicity following high-precision radiotherapy for prostate cancer. *Physics and Imaging in Radiation Oncology*, 7:65–69.

[36] Casares-Magaz, O., Thor, M., Liao, D., Frøkjær, J. B., Kræmer, P., Krogh, K., Drewes, A. M., Gregersen, H., Moiseenko, V., Høyer, M., and Muren, L. P. (2015). An image-based method to quantify biomechanical properties of the rectum in radiotherapy of prostate cancer. *Acta oncologica (Stockholm, Sweden)*, 54(9):1335–42.

[37] Catton, C. N., Gu, C.-S., Martin, J. M., Supiot, S., Bauman, G. S., Dayes, I. S., Lukka, H., Bahary, J.-P., Julian, J. A., Wu, J. S., Corbett, T. B., Tsakiridis, T., Ahmed, S., Parliament, M. B., Chung, P. W., Warde, P., Levine, M. N., Tai, K. H., Tang, C., Cheung, P., and Craig, T. K. (2017). Randomized Trial of a Hypofractionated Radiation Regimen for the Treatment of Localized Prostate Cancer. *Journal of Clinical Oncology*, 35(17):1884–1890.

[38] Cazoulat, G., Owen, D., Matuszak, M. M., Balter, J. M., and Brock, K. K. (2016). Biomechanical deformable image registration of longitudinal lung CT images using vessel information. *Physics in medicine and biology*, 61(13):4826–39.

[39] Chan, T. F. and Vese, L. A. (2001). Active contours without edges. *IEEE Transactions on Image Processing*, 10(2):266–277.

[40] Chao, M., Xie, Y., and Xing, L. (2008). Auto-propagation of contours for adaptive prostate radiation therapy. *Physics in Medicine and Biology*, 53(17):4533–4542.

[41] Chen, B., Acosta, O., Kachenoura, A., Ospina, J. D., Dréan, G., Simon, A., Bellanger, J.-J., Haigron, P., and de Crevoisier, R. (2011). *Spatial Characterization and Classification of Rectal Bleeding in Prostate Cancer Radiotherapy with a Voxel-Based Principal Components Analysis Model for 3D Dose Distribution*, pages 60–69. Springer Berlin Heidelberg, Berlin, Heidelberg.

[42] Chen, J., Chen, H., Zhong, Z., Wang, Z., Hrycushko, B., Zhou, L., Jiang, S., Albuquerque, K., Gu, X., and Zhen, X. (2018). Investigating rectal toxicity associated dosimetric features with deformable accumulated rectal surface dose maps for cervical cancer radiotherapy. *Radiation Oncology*, 13(1):125.

[43] Chen, L., Paskalev, K., Xu, X., Zhu, J., Wang, L., Price, R. A., Hu, W., Feigenberg, S. J., Horwitz, E. M., Pollack, A., and Charlie Ma, C. (2010). Rectal dose variation during the course of image-guided radiation therapy of prostate cancer. *Radiotherapy and Oncology*, 95(2):198–202.

[44] Collery, A. and Forde, E. (2017). Daily Rectal Dose-volume Histogram Variation in Prostate Intensity-modulated Radiation Therapy: Is It Clinically Significant in the Era of Image Guidance? *Journal of Medical Imaging and Radiation Sciences*, 48(4):346–351.

[45] Collins, G. S., Reitsma, J. B., Altman, D. G., and Moons, K. G. M. (2015). Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): The TRIPOD Statement. *Ann Intern Med*, 162(1):55–63.

[46] CompRT VoxTox (2016). Linking radiation dose at the voxel level with toxicity. http://www.comprt.org/research/voxtox. Accessed: August 2016.

[47] Cox, J. D., Stetz, J., and Pajak, T. F. (1995). Toxicity criteria of the Radiation Therapy Oncology Group (RTOG) and the European organization for research and treatment of cancer (EORTC). *International Journal of Radiation Oncology*Biology*Physics*, 31(5):1341–1346.

[48] Creak, A., Hall, E., Horwich, A., Eeles, R., Khoo, V., Huddart, R., Parker, C., Griffin, C., Bidmead, M., Warrington, J., and Dearnaley, D. (2013). Randomised pilot study of dose escalation using conformal radiotherapy in prostate cancer: long-term follow-up. *British Journal of Cancer*, 109.

[49] de Crevoisier, R., Fiorino, C., and Dubray, B. (2010). [Dosimetric factors predictive of late toxicity in prostate cancer radiotherapy]. *Cancer radiothérapie : journal de la Société française de radiothérapie oncologique*, 14(6-7):460–8.

[50] Dearnaley, D., Jovic, G., Syndikus, I., Khoo, V., Cowan, R. A., Graham, J. D., Aird, E. G., Bottomley, D., Huddart, R. A., Jose, C. C., Matthews, J. H. L., Millar, J. L., Murphy, C., Russell, J. M., Scrase, C. D., Parmar, M. K. B., and Sydes, M. R. (2014a). Escalated-dose versus control-dose conformal radiotherapy for prostate cancer: long-term results from the MRC RT01 randomised controlled trial. *The Lancet. Oncology*, 15(4):464–73.

[51] Dearnaley, D., Jovic, G., Syndikus, I., Khoo, V., Cowan, R. A., Graham, J. D., Aird, E. G., Bottomley, D., Huddart, R. A., Jose, C. C., Matthews, J. H. L., Millar, J. L., Murphy, C., Russell, J. M., Scrase, C. D., Parmar, M. K. B., and Sydes, M. R. (2014b). Escalated-dose versus control-dose conformal radiotherapy for prostate cancer: long-term results from the MRC RT01 randomised controlled trial. *The Lancet Oncology*, 15(4):464–473.

[52] Dearnaley, D., Syndikus, I., Mossop, H., Khoo, V., Birtle, A., Bloomfield, D., Graham, J., Kirkbride, P., Logue, J., Malik, Z., Money-Kyrle, J., O'Sullivan, J. M., Panades, M., Parker, C., Patterson, H., Scrase, C., Staffurth, J., Stockdale, A., Tremlett, J., Bidmead, M., Mayles, H., Naismith, O., South, C., Gao, A., Cruickshank, C., Hassan, S., Pugh, J., Griffin, C., and Hall, E. (2016). Conventional versus hypofractionated high-dose intensity-modulated radiotherapy for prostate cancer: 5-year outcomes of the randomised, non-inferiority, phase 3 {CHHiP} trial. *The Lancet Oncology*, 17(8):1047 – 1060.

[53] Dearnaley, D., Syndikus, I., Sumo, G., Bidmead, A. M., Bloomfield, D., Clark, C., Gao, A., Hassan, S., Horwich, A., Huddart, R., Khoo, V., Kirkbride, P., Mayles, H., Mayles, P., Naismith, O., Parker, C., Patterson, H., Russell, M., Scrase, C., South, C., Staffurth, J., and Hall, E. (2012a). Conventional versus hypofractionated high-dose intensity-modulated radiotherapy for prostate cancer: preliminary safety results from the CHHiP randomised controlled trial. *The Lancet. Oncology*, 13(1):43–54.

[54] Dearnaley, D., Syndikus, I., Sumo, G., Bidmead, M., Bloomfield, D., Clark, C., Gao, A., Hassan, S., Horwich, A., Huddart, R., Khoo, V., Kirkbride, P., Mayles, H., Mayles, P., Naismith, O., Parker, C., Patterson, H., Russell, M., Scrase, C., South, C., Staffurth, J., and Hall, E. (2012b). Conventional versus hypofractionated high-dose intensity-modulated radiotherapy for prostate cancer: preliminary safety results from the {CHHiP} randomised controlled trial. *The Lancet Oncology*, 13(1):43 – 54.

[55] Dearnaley, D. P., Jovic, G., Syndikus, I., Khoo, V., Cowan, R. A., Graham, J. D., Aird, E. G., Bottomley, D., Huddart, R. A., Jose, C. C., Matthews, J. H. L., Millar, J. L., Murphy, C., Russell, J. M., Scrase, C. D., Parmar, M. K. B., and Sydes, M. R. (2014c). Escalated-dose versus control-dose conformal radiotherapy for prostate cancer: long-term results from the {MRC} {RT01} randomised controlled trial. *The Lancet Oncology*, 15(4):464 – 473.

[56] Dearnaley, D. P., Sydes, M. R., Graham, J. D., Aird, E. G., Bottomley, D., Cowan, R. A., Huddart, R. A., Jose, C. C., Matthews, J. H., Millar, J., Moore, A. R., Morgan, R. C., Russell, J. M., Scrase, C. D., Stephens, R. J., Syndikus, I., and Parmar, M. K. (2007a). Escalated-dose versus standard-dose conformal radiotherapy in prostate cancer: first results from the {MRC} {RT01} randomised controlled trial. *The Lancet Oncology*, 8(6):475 – 487.

[57] Dearnaley, D. P., Sydes, M. R., Langley, R. E., Graham, J. D., Huddart, R. A., Syndikus, I., Matthews, J. H., Scrase, C. D., Jose, C. C., Logue, J., and Stephens, R. J. (2007b). The early toxicity of escalated versus standard dose conformal radiotherapy with neo-adjuvant androgen suppression for patients with localised prostate cancer: Results from the MRC RT01 trial (ISRCTN47772397) . *Radiotherapy and Oncology*, 83(1):31 – 41.

[58] Deore, S. M., Shrivastava, S. K., Supe, S. J., Viswanathan, P. S., and Dinshaw, K. A. (1993). Alpha/beta value and importance of dose per fraction for the late rectal and recto-sigmoid complications. *Strahlentherapie und Onkologie*, 169(9):521–526.

[59] Dréan, G., Acosta, O., Lafond, C., Simon, A., de Crevoisier, R., and Haigron, P. (2016a). Inter-individual registration and dose mapping for voxelwise population analysis of rectal toxicity in prostate cancer radiotherapy. *Medical Physics*, 43(6):2721–2730.

[60] Dréan, G., Acosta, O., Ospina, J. D., Fargeas, A., Lafond, C., Corrégé, G., Lagrange, J.-L., Créhange, G., Simon, A., Haigron, P., and de Crevoisier, R. (2016b). Identification of a rectal subregion highly predictive of rectal bleeding in prostate cancer IMRT. *Radiotherapy and Oncology*, 119(3):388–397.

[61] Dréan, G., Acosta, O., Ospina, J. D., Fargeas, A., Lafond, C., Corrégé, G., Lagrange, J.-L., Créhange, G., Simon, A., Haigron, P., and de Crevoisier, R. (2016). Identification of a rectal subregion highly predictive of rectal bleeding in prostate cancer IMRT. *Radiotherapy and Oncology*, 119(3):388 – 397.

[62] Eckert, F., Alloussi, S., Paulsen, F., Bamberg, M., Zips, D., Spillner, P., Gani, C., Kramer, U., Thorwarth, D., Schilling, D., and Müller, A.-C. (2013). Prospective evaluation of a hydrogel spacer for rectal separation in dose-escalated intensity-modulated radiotherapy for clinically localized prostate cancer. *BMC cancer*, 13:27.

[63] Emami, B., Lyman, J., Brown, A., Cola, L., Goitein, M., Munzenrider, J., Shank, B., Solin, L., and Wesson, M. (1991). Tolerance of normal tissue to therapeutic irradiation. *International Journal of Radiation Oncology*Biology*Physics*, 21(1):109 – 122.

[64] Fargeas, A., Acosta, O., David, J., Arrango, O., Ferhat, A., Costet, N., Albera, L., Azria, D., Fenoglietto, P., Créhange, G., Beckendorf, V., Hatt, M., Kachenoura, A., and De Crevoisier, R. (2017). Independent component analysis for rectal bleeding prediction following prostate cancer radiotherapy. *Radiotherapy and Oncology*, 119(3):388–397.

[65] Fung, Y. (1993). *Biomechanics; Mechanical Properties of Living Tissues*. Springer-Verlag New York.

[66] Gay, H. A. and Niemierko, A. (2007). A free program for calculating EUD-based NTCP and TCP in external beam radiotherapy. *Physica Medica*, 23(3-4):115–125.

[67] Geraghty, J. P., Grogan, G., and Ebert, M. A. (2013). Automatic segmentation of male pelvic anatomy on computed tomography images: A comparison with multiple observers in the context of a multicentre clinical trial. *Radiation Oncology*, 8(1):106 – 116.

[68] Gilles Marckmann, E. V. (2006). Comparison of hyperelastic models for rubber-like materials. Rubber Chemistry and Technology. *American Chemical Society*, 79(5):835–858.

[69] Glickman, M. E., Rao, S. R., and Schultz, M. R. (2014). False discovery rate control is a recommended alternative to Bonferroni-type adjustments in health studies. *Journal of Clinical Epidemiology*, 67:850–857.

[70] Gulliford, S. L., Foo, K., Morgan, R. C., Aird, E. G., Bidmead, A. M., Critchley, H., Evans, P. M., Gianolini, S., Mayles, W. P., Moore, A. R., Sánchez-Nieto, B., Partridge, M., Sydes, M. R., Webb, S., and Dearnaley, D. P. (2010a). Dose–volume constraints to reduce rectal side effects from prostate radiotherapy: Evidence from {MRC} {RT01} trial {ISRCTN} 47772397. *International Journal of Radiation Oncology*Biology*Physics*, 76(3):747 – 754.

[71] Gulliford, S. L., Partridge, M., Sydes, M. R., Andreyev, J., and Dearnaley, D. P. (2010b). A comparison of dose–volume constraints derived using peak and longitudinal definitions of late rectal toxicity. *Radiotherapy and Oncology*, 94(2):241 – 247. Selected papers from the 10th Biennial {ESTRO} Conference on Physics and Radiation Technology for Clinical Radiotherapy.

[72] Hamdy, F. C., Donovan, J. L., Lane, J. A., Mason, M., Metcalfe, C., Holding, P., Davis, M., Peters, T. J., Turner, E. L., Martin, R. M., Oxley, J., Robinson, M., Staffurth, J., Walsh, E., Bollina, P., Catto, J., Doble, A., Doherty, A., Gillatt, D., Kockelbergh, R., Kynaston, H., Paul, A., Powell, P., Prescott, S., Rosario, D. J., Rowe, E., and Neal, D. E. (2016). 10-Year Outcomes after Monitoring, Surgery, or Radiotherapy for Localized Prostate Cancer Supplementary Appendix. *New England Journal of Medicine*, 375(15):1415–1424.

[73] Hamlett, L. J., McPartlin, A. J., Maile, E. J., Webster, G., Swindell, R., Rowbottom, C. G., Choudhury, A., and Aitkenhead, A. H. (2015a). Parametrized rectal dose and associations with late toxicity in prostate cancer radiotherapy. *The British Journal of Radiology*, 88(1054):20150110. PMID: 26246172.

[74] Hamlett, L. J., McPartlin, A. J., Maile, E. J., Webster, G., Swindell, R., Rowbottom, C. G., Choudhury, A., and Aitkenhead, A. H. (2015b). Parametrized rectal dose and associations with late toxicity in prostate cancer radiotherapy. *The British Journal of Radiology*, 88(1054):20150110.

[75] Hatton, J. A., Greer, P. B., Tang, C., Wright, P., Capp, A., Gupta, S., Parker, J., Wratten, C., and Denham, J. W. (2011). Does the planning dose–volume histogram represent treatment doses in image-guided prostate radiation therapy? Assessment with cone-beam computerised tomography scans. *Radiotherapy and Oncology*, 98(2):162–168.

[76] Hensel, J. M., Ménard, C., Chung, P. W., Milosevic, M. F., Kirilova, A., Moseley, J. L., Haider, M. A., and Brock, K. K. (2007). Development of Multiorgan Finite Element-Based Prostate Deformation Model Enabling Registration of Endorectal Coil Magnetic Resonance Imaging for Radiotherapy Planning. *International Journal of Radiation Oncology*Biology*Physics*, 68(5):1522–1528.

[77] Holch, P., Henry, A. M., Davidson, S., Gilbert, A., Routledge, J., Shearsmith, L., Franks, K., Ingleson, E., Albutt, A., and Velikova, G. (2017). Acute and Late Adverse Events Associated With Radical Radiation Therapy Prostate Cancer Treatment: A Systematic Review of Clinician and Patient Toxicity Reporting in Randomized Controlled Trials. *Int J Radiation Oncol Biol Phys*, 97(3):495–510.

[78] Holyoake, D. L. P., Aznar, M., Mukherjee, S., Partridge, M., and Hawkins, M. A. (2017). Modelling duodenum radiotherapy toxicity using cohort dose-volume-histogram data. *Radiotherapy and Oncology*, 123:431–437.

[79] Huang, Y., Joiner, M., Zhao, B., Liao, Y., and Burmeister, J. (2010). Dose convolution filter: Incorporating spatial dose information into tissue response modeling. *Medical Physics*, 37(3):1068–1074.

[80] Hutchinson, R. C., Sundaram, V., Folkert, M., and Lotan, Y. (2016). Decision analysis model evaluating the cost of a temporary hydrogel rectal spacer before prostate radiation therapy to reduce the incidence of rectal complications. *Urologic oncology*, 34(7):291.e19–26.

[81] ICRU (1999). Report 62: Prescribing, Recording and Reporting Photon Beam Therapy (Supplement to ICRU Report 50). *International Commission on Radiation Units and Measurements*.

[82] Jaffray, D. A., Lindsay, P. E., Brock, K. K., Deasy, J. O., and Tomé, W. (2010). Accurate Accumulation of Dose for Improved Understanding of Radiation Effects in Normal Tissue. *International Journal of Radiation Oncology*Biology*Physics*, 76(3):S135–S139.

[83] Keros, L., Bernier, V., Aletti, P., Marchesi, V., Wolf, D., and Noel, A. (2006). Qualitative estimation of pelvic organ interactions and their consequences on prostate motion: Study on a deceased person. *Medical Physics*, 33(6):1902–1910.

[84] Kuban, D. A., Tucker, S. L., Dong, L., Starkschall, G., Huang, E. H., Cheung, M. R., Lee, A. K., and Pollack, A. (2008). Long-Term Results of the M. D. Anderson Randomized Dose-Escalation Trial for Prostate Cancer. *International Journal of Radiation Oncology*Biology*Physics*, 70(1):67–74.

[85] Kupelian, P. A., Langen, K. M., Zeidan, O. A., Meeks, S. L., Willoughby, T. R., Wagner, T. H., Jeswani, S., Ruchala, K. J., Haimerl, J., and Olivera, G. H. (2006). Daily variations in delivered doses in patients treated with radiotherapy for localized prostate cancer. *International Journal of Radiation Oncology*Biology*Physics*, 66(3):876–882.

[86] Kutcher, G., Burman, C., Brewster, L., Goitein, M., and Mohan, R. (1991). Histogram reduction method for calculating complication probabilities for three-dimensional treatment planning evaluations. *International Journal of Radiation Oncology*Biology*Physics*, 21(1):137 – 146.

[87] Landoni, V., Fiorino, C., Cozzarini, C., Sanguineti, G., Valdagni, R., and Rancati, T. (2016). Predicting toxicity in radiotherapy for prostate cancer. *Physica Medica*, 32(3):521 – 532.

[88] Lee, W. R., Dignam, J. J., Amin, M., Bruner, D., Low, D., Swanson, G. P., Shah, A., D'Souza, D., Michalski, J. M., Dayes, I., Seaward, S. A., Hall, W. A., Nguyen, P. L., Pisansky, T. M., Faria, S., Chen, Y., Koontz, B. F., Paulus, R., and Sandler, H. M. (2017). NRG Oncology RTOG 0415: A randomized phase III non-inferiority study comparing two fractionation schedules in patients with low-risk prostate cancer. *Journal of Clinical Oncology*, 34(2suppl):1–1.

[89] Litwin, M. S., Hays, R. D., Fink, A., Ganz, P. A., Leake, B., and Brook, R. H. (1998). The UCLA Prostate Cancer Index: Development, Reliability, and Validity of a Health-Related Quality of Life Measure. *Medical Care*, 36(7):1002–1012.

[90] Lütgendorf-Caucig, C., Fotina, I., Stock, M., Pötter, R., Goldner, G., and Georg, D. (2011). Feasibility of CBCT-based target and normal structure delineation in prostate cancer radiotherapy: Multi-observer and image multi-modality study. *Radiotherapy and Oncology*, 98(2):154–161.

[91] Lyman, J. T. (1985). Complication probability as assessed from dose-volume histograms. *Radiation Research*, 104(2):S13–S19.

[92] Marcello, M., Ebert, M., Haworth, A., Steigler, A., Kennedy, A., Joseph, D., and Denham, J. (2018). Association between treatment planning and delivery factors and disease progression in prostate cancer radiotherapy: Results from the TROG 03.04 RADAR trial. *Radiotherapy and Oncology*.

[93] Marchal, M., Promayon, E., and Troccaz, J. (2006). Simulating prostate surgical procedures with a discrete soft tissue model. In Mendoza, C. and Navazo, I., editors, *3rd Workshop in Virtual Reality Interactions and Physical Simulation "VRIPHYS"*. Eurographics Digital Library.

[94] Marks, L. B., Haken, R. K. T., and Martel, M. K. (2010). Guest editor's introduction to quantec: A users guide. *International Journal of Radiation Oncology*Biology*Physics*, 76(3, Supplement):S1 – S2. Quantitative Analyses of Normal Tissue Effects in the Clinic.

[95] Marzi, S., Saracino, B., Petrongari, M. G., Arcangeli, S., Gomellini, S., Arcangeli, G., Benassi, M., and Landoni, V. (2009). Modeling of α/β for late rectal toxicity from a randomized phase II study: conventional versus hypofractionated scheme for localized prostate cancer. *Journal of Experimental & Clinical Cancer Research*, 28(28).

[96] Mcparland, N., Pearson, M., Wong, J., Sigur, I., Stenger, C., and Tyldesley, S. (2018). Quantifying daily variation in volume and dose to the prostate, rectum and bladder using cone-beam computerised tomography. *Journal of Radiotherapy in Practice*, 13:79–86.

[97] Michalski, J. M., Gay, H., Jackson, A., Tucker, S. L., and Deasy, J. O. (2010). Radiation Dose-Volume Effects in Radiation-Induced Rectal Injury. *International Journal of Radiation Oncology Biology Physics*, 76(3 SUPPL.):123–129.

[98] Michalski, J. M., Moughan, J., Purdy, J., Bosch, W., Bahary, J.-P., Lau, H. Y., Duclos, M., Parliament, M., Morton, G., Hamstra, D. A., Seider, M. J., Lock, M., Patel, M., Gay, H. A., Vigneault, E., Dignam, J., and Sandler, H. M. (2017). A randomized trial of 79.2Gy versus 70.2Gy radiation therapy (RT) for localized prostate cancer. *Journal of Clinical Oncology*, 33(7suppl):4–4.

[99] Michalski, J. M., Moughan, J., Purdy, J., Bosch, W., Bruner, D. W., Bahary, J. P., Lau, H., Duclos, M., Parliament, M., Morton, G., Hamstra, D., Seider, M., Lock, M. I., Patel, M., Gay, H., Vigneault, E., Winter, K., and Sandler, H. (2018). Effect of standard vs dose-escalated radiation therapy for patients with intermediate-risk prostate cancer the nrg oncology RTOG 0126 randomized clinical trial. *JAMA Oncology*, 4(6):e180039.

[100] Mistry, M., Parkin, D. M., Ahmad, A. S., and Sasieni, P. (2011). Cancer incidence in the united kingdom: projections to the year 2030. *Br J Cancer*, 105:1795–803.

[101] Morris, K. A. and Haboubi, N. Y. (2015). Pelvic radiation therapy: Between delight and disaster. *World J Gastrointest Surg*, 7(11):279–288.

[102] Mościcki, J. T., Brochu, F., Ebke, J., Egede, U., Elmsheuser, J., Harrison, K., Jones, R. W. L., Lee, H. C., Liko, D., Maier, A., Muraru, A., Patrick, G. N., Pajchel, K., Reece, W., Samset, B. H., Slater, M. W., Soroko, A., Tan, C. L., Vanderster, D. C., and Williams, M. (2009). Ganga: a tool for computational-task management and easy access to Grid resources. 180(11):2303–2316.

[103] Moulton, C. R., House, M. J., Lye, V., Tang, C. I., Krawiec, M., Joseph, D. J., Denham, J. W., and Ebert, M. A. (2017). Spatial features of dose–surface maps from deformably-registered plans correlate with late gastrointestinal complications. *Phys. Med. Biol*, 62(10):4118–4139.

[104] Murray, J., McQuaid, D., Dunlop, A., Buettner, F., Nill, S., Hall, E., Dearnaley, D., and Gulliford, S. (2014). Su-e-j-14: A novel approach to evaluate the dosimetric effect of rectal variation during image guided prostate radiotherapy. *Medical Physics*, 41(6):157.

[105] Mylona, E., Acosta, O., Lizee, T., Lafond, C., Crehange, G., Magné, N., Chiavassa, S., Supiot, S., Ospina Arango, J. D., Campillo-Gimenez, B., Castelli, J., and de Crevoisier, R. (2019). Voxel-Based Analysis for Identification of Urethrovesical Subregions Predicting Urinary Toxicity After Prostate Cancer Radiation Therapy. *International Journal of Radiation Oncology Biology Physics*, 104(2):343–354.

[106] National Cancer Registration and Analysis Service, Cancer Research UK, and Public Health England (2018). Chemotherapy, radiotherapy and tumour resection by cancer alliance in england, 2013 - 2015, workbook. http://www.ncin.org.uk/cancer_type_and_topic_specific_work/topic_specific_work/main_cancer_treatments. Accessed: July 2018.

[107] National Institute for Health and Care Excellence (2014). Prostate cancer: diagnosis and treatment, clinical guideline.

[108] National Institute for Health and Care Excellence (2016). Exceptional surveillance review 2016 – Prostate cancer: diagnosis and management (2014) NICE guideline CG175, Appendix A: summary of new evidence.

[109] National Institute for Health and Care Excellenece (NICE) (2017). Biodegradable spacer insertion to reduce rectal toxicity during radiotherapy for prostate cancer (ipg590). https://www.nice.org.uk/guidance/ipg590. Accessed: May 2019.

[110] Ng, A., Brock, K. K., Sharpe, M. B., Moseley, J. L., Craig, T., and Hodgson, D. C. (2012). Individualized 3d reconstruction of normal tissue dose for patients with long-term follow-up: A step toward understanding dose risk for late toxicity. *International Journal of Radiation Oncology*Biology*Physics*, 84(4):e557 – e563.

[111] Ng, A., Nguyen, T. N., Moseley, J. L., Hodgson, D. C., Sharpe, M. B., and Brock, K. K. (2010). Reconstruction of 3D lung models from 2D planning data sets for Hodgkin's lymphoma patients using combined deformable image registration and navigator channels. *Med Phys*, 37(3):1017–1028.

[112] Ng, M., Brown, E., Williams, A., Chao, M., Lawrentschuk, N., and Chee, R. (2014). Fiducial markers and spacers in prostate radiotherapy: current applications. *BJU international*, 113 Suppl:13–20.

[113] Niemierko, A. (1997). Reporting and analyzing dose distributions: A concept of equivalent uniform dose. *Medical Physics*, 24(1):103–110.

[114] Niu, C. J., Foltz, W. D., Velec, M., Moseley, J. L., Al-Mayah, A., and Brock, K. K. (2012). A novel technique to enable experimental validation of deformable dose accumulation. *Medical Physics*, 39(2):765–776.

[115] Onjukka, E., Fiorino, C., Cicchetti, A., Palorini, F., Improta, I., Gagliardi, G., Cozzarini, C., Degli Esposti, C., Gabriele, P., Valdagni, R., and Rancati, T. (2019). Patterns in ano-rectal dose maps and the risk of late toxicity after prostate IMRT. *Acta Oncologica*, pages 1–8.

[116] Ospina, J. D., Commandeur, F., Ríos, R., Dréan, G., Correa, J. C., Simon, A., Haigron, P., de Crevoisier, R., and Acosta, O. (2013). A tensor-based population value decomposition to explain rectal toxicity after prostate cancer radiotherapy. *Medical image computing and computer-assisted intervention: MICCAI ... International Conference on Medical Image Computing and Computer-Assisted Intervention*, 16(Pt 2):387–394.

[117] Padhani, A. R., Khoo, V. S., Suckling, J., Husband, J. E., Leach, M. O., and Dearnaley, D. (1999). Evaluating the effect of rectal distension and rectal movement on prostate gland position using cine MRI. *International Journal of Radiation Oncology*Biology*Physics*, 44(3):525–533.

[118] Palma, G., Monti, S., D'Avino, V., Conson, M., Liuzzi, R., Pressello, M. C., Donato, V., Deasy, J. O., Quarantelli, M., Pacelli, R., and Cella, L. (2016). A Voxel-Based Approach to Explore Local Dose Differences Associated With Radiation-Induced Lung Damage. *International Journal of Radiation Oncology Biology Physics*, 96(1):127–133.

[119] Palma, G., Monti, S., Thor, M., Soscia, E., Sirignano, C., Rimner, A., Deasy, J., and Cella, L. (2018). Redesign of Voxel-Based Analysis for SBRT Lung Cancer Patients and Refinement of Findings on Regional Dose Differences Associated with Radiation-Induced Acute Lung Damage. *International Journal of Radiation Oncology*Biology*Physics*, 102(3S):S95.

[120] Palma, G., Monti, S., Xu, T., Scifoni, E., Yang, P., Hahn, S. M., Durante, M., Mohan, R., Liao, Z., and Cella, L. (2019). Spatial Dose Patterns Associated With Radiation Pneumonitis in a Randomized Trial Comparing Intensity-Modulated Photon Therapy With Passive Scattering Proton Therapy for Locally Advanced Non-Small Cell Lung Cancer. *International journal of radiation oncology, biology, physics*, 104(5):1124–1132.

[121] Palorini, F., Botti, A., Carillo, V., Gianolini, S., Improta, I., Iotti, C., Rancati, T., Cozzarini, C., and Fiorino, C. (2016a). Bladder dose-–surface maps and urinary toxicity: Robustness with respect to motion in assessing local dose effects. *Physica Medica*, 32(3):506–511.

[122] Palorini, F., Cozzarini, C., Gianolini, S., Botti, A., Carillo, V., Iotti, C., Rancati, T., Valdagni, R., and Fiorino, C. (2016b). First application of a pixel-wise analysis on bladder dose–surface maps in prostate cancer radiotherapy. *Radiotherapy and Oncology*, 119(1):123 – 128.

[123] Pearson, D., Gill, S. K., Campbell, N., and Reddy, K. (2016). Dosimetric and volumetric changes in the rectum and bladder in patients receiving CBCT-guided prostate IMRT: analysis based on daily CBCT dose calculation. *Journal of Applied Clinical Medical Physics*, 17(6):107–117.

[124] Peeters, S. T. H., Lebesque, J. V., Heemsbergen, W. D., Van Putten, W. L. J., Slot, A., Dielwart, M. F. H., and Koper, P. C. M. (2006). Localized volume effects for late rectal and anal toxicity after radiotherapy for prostate cancer. *International Journal of Radiation Oncology Biology Physics*, 64(4):1151–1161.

[125] Peng, C., Ahunbay, E., Chen, G., Anderson, S., Lawton, C., and Li, X. A. (2011). Characterizing Interfraction Variations and Their Dosimetric Effects in Prostate Cancer Radiotherapy. *International Journal of Radiation Oncology*Biology*Physics*, 79(3):909–914.

[126] Peterson, J. L., Buskirk, S. J., Heckman, M. G., Diehl, N. N., Bernard, J. R., Tzou, K. S., Casale, H. E., Bellefontaine, L. P., Serago, C., Kim, S., Vallow, L. a., Daugherty, L. C., and Ko, S. J. (2014). Image-guided intensity-modulated radiotherapy for prostate cancer: Dose constraints for the anterior rectal wall to minimize rectal toxicity. *Medical Dosimetry*, 39(1):12–17.

[127] Pilepich, M., Asbell, S., Krall, J., Baerwald, W., Sause, W., Rubin, P., Emami, B., and Pidcock, G. (1987). Correlation of radiotherapeutic parameters and treatment related morbidity—analysis of rtog study 77-06. *International Journal of Radiation Oncology*Biology*Physics*, 13(7):1007 – 1012.

[128] Pinkawa, M., Piroth, M. D., Holy, R., Escobar-Corral, N., Caffaro, M., Djukic, V., Klotz, J., and Eble, M. J. (2013). Spacer stability and prostate position variability during radiotherapy for prostate cancer applying a hydrogel to protect the rectal wall. *Radiotherapy and Oncology*, 106(2):220–224.

[129] Pollack, A., Walker, G., Ma, C., Stoyanova, R., Feigenberg, S., Horwitz, E. M., Buyyounouski, M. K., Movsas, B., Price, R., Uzzo, R. G., Greenberg, R. E., and Konski, A. A. (2013). Randomized Trial of Hypofractionated External-Beam Radiotherapy for Prostate Cancer. *Journal of Clinical Oncology*, 31(31):3860–3868.

[130] Pollack, A., Zagars, G. K., Starkschall, G., Antolak, J. A., Lee, J. J., Huang, E., Von Eschenbach, A. C., Kuban, D. A., and Rosen, I. (2002). Prostate cancer radiation dose response: Results of the M. D. Anderson phase III randomized trial. *International Journal of Radiation Oncology Biology Physics*, 53(5):1097–1105.

[131] Rancati, T., Fiorino, C., Gagliardi, G., Cattaneo, G. M., Sanguineti, G., Borca, V. C., Cozzarini, C., Fellin, G., Foppiano, F., Girelli, G., Menegotti, L., Piazzolla, A., Vavassori, V., and Valdagni, R. (2004). Fitting late rectal bleeding data using different NTCP models: Results from an Italian multi-centric study (AIROPROS0101). *Radiotherapy and Oncology*, 73(1):21–32.

[132] Romanchikova, M., Harrison, K., Burnet, N. G., Hoole, A. F., Sutcliffe, M. P., Parker, M. A., Jena, R., and Thomas, S. J. (2017). Automated retrieval of radiotherapy data for clinical trials, audit and research. *The British Journal of Radiology.*, 90:20170651.

[133] Royal College of Radiologists (2019). Radiotherapy dose fractionation, third edition - prostate cancer.

[134] Samavati, N., McGrath, D. M., Jewett, M. A. S., van der Kwast, T., Ménard, C., and Brock, K. K. (2015a). Effect of material property heterogeneity on biomechanical modeling of prostate under deformation. *Physics in Medicine and Biology*, 60(1):195.

[135] Samavati, N., Velec, M., and Brock, K. K. (2015b). A hybrid biomechanical intensity based deformable image registration of lung 4DCT. *Physics in medicine and biology*, 60(8):3359–73.

[136] Scaife, J., Harrison, K., Romanchikova, M., Parker, A., Sutcliffe, M., Bond, S., Thomas, S., Freeman, S., Jena, R., Bates, A., and Burnet, N. (2014). Random variation in rectal position during radiotherapy for prostate cancer is two to three times greater than that predicted from prostate motion. *The British Journal of Radiology*, 87(1042):20140343. PMID: 25138155.

[137] Scaife, J. E. (2016). Calculating accumulated radiotherapy dose to the rectum in order to improve late toxicity prediction in patients treated for prostate cancer. University of Cambridge, Department of Oncology, PhD thesis 39445.

[138] Scaife, J. E., Harrison, K., Drew, A., Cai, X., Lee, J., Schonlieb, C.-B., Sutcliffe, M., Parker, M. A., Freeman, S., Romanchikova, M., Thomas, S. J., Jena, R., Bates, A., and Burnet, N. (2015a). Accuracy of manual and automated rectal contours using helical tomotherapy image guidance scans during prostate radiotherapy. *Journal of Clinical Oncology*, 33(7_suppl):94–94.

[139] Scaife, J. E., Thomas, S. J., Harrison, K., Romanchikova, M., Sutcliffe, M. P. F., Forman, J. R., Bates, A. M., Jena, R., Parker, M. A., and Burnet, N. G. (2015b). Accumulated dose to the rectum, measured using dose-volume histograms and dose-surface maps, is different from planned dose in all patients treated with radiotherapy for prostate cancer. *The British Journal of Radiology*, 88(1054):20150243.

[140] Schaake, W., van der Schaaf, A., van Dijk, L. V., Bongaerts, A. H., van den Bergh, A. C., and Langendijk, J. A. (2016). Normal tissue complication probability (ntcp) models for late rectal bleeding, stool frequency and fecal incontinence after radiotherapy in prostate cancer patients. *Radiotherapy and Oncology*, 119(3):381 – 387.

[141] Serrano, N. A., Kalman, N. S., and Anscher, M. S. (2017). Reducing rectal injury in men receiving prostate cancer radiation therapy: Current perspectives. 2017(9):339–350.

[142] Shelley, L. E. A., Scaife, J. E., Romanchikova, M., Harrison, K., Forman, J. R., Bates, A. M., Noble, D. J., Jena, R., Parker, M. A., Sutcliffe, M. P. F., Thomas, S. J., and Burnet, N. G. (2017). Delivered dose can be a better predictor of rectal toxicity than planned dose in prostate radiotherapy. *Radiotherapy and Oncology*, 123:466–471.

[143] Shelley, L. E. A., Sutcliffe, M. P. F., Harrison, K., Scaife, J. E., Parker, M. A., Romanchikova, M., Thomas, S. J., Jena, R., and Burnet, N. G. (2019). Autosegmentation of the rectum on megavoltage image guidance scans. *Biomedical Physics and Engineering Express*, 5(2):025006.

[144] Sonke, J.-J., Aznar, M., and Rasch, C. (2019). Adaptive Radiotherapy for Anatomical Changes. *Seminars in Radiation Oncology*, 29(3):245–257.

[145] Sripadam, R., Stratford, J., Henry, A. M., Jackson, A., Moore, C. J., and Price, P. (2009). Rectal motion can reduce CTV coverage and increase rectal dose during prostate radiotherapy: A daily cone-beam CT study. *Radiotherapy and Oncology*, 90:312–317.

[146] Strimmer, K. (2008). A unified approach to false discovery rate estimation. *BMC Bioinformatics*, 9(303).

[147] Susil, R. C., McNutt, T. R., DeWeese, T. L., and Song, D. (2010). Effects of Prostate-Rectum Separation on Rectal Dose From External Beam Radiotherapy. *International Journal of Radiation Oncology Biology Physics*, 76(4):1251–1258.

[148] Sydes, M. R., Stephens, R. J., Moore, A., Aird, E. G., Bidmead, A., Fallowfield, L. J., Graham, J., Griffiths, S., Mayles, W., McGuire, A., Stanley, S., Warrington, A., the RT01 collaborators, and Dearnaley, D. P. (2004). Implementing the {UK} medical research council (mrc) {RT01} trial (isrctn 47772397): methods and practicalities of a randomised controlled trial of conformal radiotherapy in men with localised prostate cancer. *Radiotherapy and Oncology*, 72(2):199 – 211.

[149] Syndikus, I., Morgan, R. C., Sydes, M. R., Graham, J. D., and Dearnaley, D. P. (2010). Late gastrointestinal toxicity after dose-escalated conformal radiotherapy for early prostate cancer: Results from the {UK} medical research council {RT01} trial (isrctn47772397). *International Journal of Radiation Oncology*Biology*Physics*, 77(3):773 – 783.

[150] Söhn, M., Alber, M., and Yan, D. (2007). Principal component analysis-based pattern analysis of dose–volume histograms and influence on rectal toxicity. *International Journal of Radiation Oncology*Biology*Physics*, 69(1):230 – 239.

[151] Thirion, J.-P. (1998). Image matching as a diffusion process: an analogy with maxwell's demons. *Medical Image Analysis*, 2(3):243 – 260.

[152] Thomas, S., Eyre, K. R., Tudor, G. S. J., and Fairfoul, J. (2012). Dose calculation software for helical tomotherapy, utilizing patient CT data to calculate an independent three-dimensional dose cube. *Medical physics*, 39(1):160–7.

[153] Thomas, S. J., Ashburner, M., Tudor, G. S. J., Treeby, J., Dean, J., Routsis, D., Rimmer, Y. L., Russell, S. G., and Burnet, N. (2013). Intra-fraction motion of the prostate during treatment with helical tomotherapy. *Radiotherapy and Oncology*, 109(3):482–486.

[154] Thomas, S. J., Romanchikova, M., Harrison, K., Parker, M. A., Bates, A. M., Scaife, J. E., Sutcliffe, M. P. F., and Burnet, N. G. (2016). Recalculation of dose for each fraction of treatment on TomoTherapy. *BRITISH JOURNAL OF RADIOLOGY*, 89(1059).

[155] Thor, M., Apte, A., Deasy, J. O., and Muren, L. P. (2013a). Statistical simulations to estimate motion-inclusive dose-volume histograms for prediction of rectal morbidity following radiotherapy. *Acta Oncologica*, 52(3):666–675.

[156] Thor, M., Bentzen, L., Hysing, L. B., Ekanger, C., Helle, S.-I., Karlsdóttir, Á., and Muren, L. P. (2013b). Prediction of rectum and bladder morbidity following radiotherapy of prostate cancer based on motion-inclusive dose distributions. *Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology*, 107(2):147–52.

[157] Tijssen, R. H., Philippens, M. E., Paulson, E. S., Glitzner, M., Chugh, B., Wetscherek, A., Dubec, M., Wang, J., and van der Heide, U. A. (2019). MRI commissioning of 1.5T MR-linac systems – a multi-institutional study. *Radiotherapy and Oncology*, 132:114–120.

[158] U.S. department of health and human services NIoH (2016). National cancer institute. common terminology criteria for adverse events (ctcae) version 4.0 2010. http://evs.nci.nih.gov/ftp1/CTCAE/CTCAE_4.03_2010-06-14_QuickReference_5x7.pdf. Accessed: July 2016.

[159] Van Gysen, K., Kneebone, A., Alfieri, F., Guo, L., and Eade, T. (2014). Feasibility of and rectal dosimetry improvement with the use of SpaceOAR® hydrogel for dose-escalated prostate cancer radiotherapy. *Journal of Medical Imaging and Radiation Oncology*, 58:511–516.

[160] Vanneste, B. G., Buettner, F., Pinkawa, M., Lambin, P., and Hoffmann, A. L. (2018). Ano-rectal wall dose-surface maps localize the dosimetric benefit of hydrogel rectum spacers in prostate cancer radiotherapy. *Clinical and Translational Radiation Oncology*, 14:17–24.

[161] Varadhan, R., Magome, T., and Hui, S. (2015). Characterization of deformation and physical force in uniform low contrast anatomy and its impact on accuracy of deformable image registration. *Medical Physics*, 43(1):52–61.

[162] Velec, M., Juang, T., Moseley, J. L., Oldham, M., and Brock, K. K. (2015). Utility and validation of biomechanical deformable image registration in low-contrast images. *Practical Radiation Oncology*, 5(4):e401 – e408.

[163] Velec, M., Moseley, J. L., Craig, T., Dawson, L. A., and Brock, K. K. (2012). Accumulated Dose in Liver Stereotactic Body Radiotherapy: Positioning, Breathing, and Deformation Effects. *International Journal of Radiation Oncology*Biology*Physics*, 83(4):1132–1140.

[164] Velec, M., Moseley, J. L., Dawson, L. A., and Brock, K. K. (2014). Dose Escalated Liver Stereotactic Body Radiation Therapy at the Mean Respiratory Position. *International Journal of Radiation Oncology*Biology*Physics*, 89(5):1121–1128.

[165] Velec, M., Moseley, J. L., Svensson, S., Hårdemark, B., Jaffray, D. A., and Brock, K. K. (2017). Validation of biomechanical deformable image registration in the abdomen, thorax, and pelvis in a commercial radiotherapy treatment planning system. *Medical Physics*, 44(7):3407–3417.

[166] Viani, G. A., Stefano, E. J., and Afonso, S. L. (2009). Higher-than-conventional radiation doses in localized prostate cancer treatment: a meta-analysis of randomized, controlled trials. *International journal of radiation oncology, biology, physics*, 74(5):1405–18.

[167] Widmark, A., Gunnlaugsson, A., Beckman, L., Thellenberg-Karlsson, C., Hoyer, M., Lagerlund, M., Kindblom, J., Ginman, C., Johansson, B., Björnlinger, K., Seke, M., Agrup, M., Fransson, P., Tavelin, B., Norman, D., Zackrisson, B., Anderson, H., Kjellén, E., Franzén, L., and Nilsson, P. (2019). Ultra-hypofractionated versus conventionally fractionated radiotherapy for prostate cancer: 5-year outcomes of the HYPO-RT-PC randomised, non-inferiority, phase 3 trial. *The Lancet*, 394(10196):385–395.

[168] Wilkins, A., Naismith, O. F., Brand, D., Fernandez, K., Hall, E., Dearnaley, D. P., and Gulliford, S. (2018). Derivation of dose/volume constraints for the anorectum from clinician-reported outcomes (CRO) in the CHHiP trial of radiotherapy (RT) fractionation. *Journal of Clinical Oncology*, 36(6suppl):87.

[169] Witztum, A., George, B., Warren, S., Partridge, M., and Hawkins, M. A. (2016). Unwrapping 3D complex hollow organs for spatial dose surface analysis. *Medical Physics*, 43(11):6009–6016.

[170] Wortel, R. C., Incrocci, L., Pos, F. J., van der Heide, U. A., Lebesque, J. V., Aluwini, S., Witte, M. G., and Heemsbergen, W. D. (2016). Late Side Effects After Image Guided Intensity Modulated Radiation Therapy Compared to 3D-Conformal Radiation Therapy for Prostate Cancer: Results From 2 Prospective Cohorts. *International Journal of Radiation Oncology*Biology*Physics*, 95(2):680–689.

[171] Wortel, R. C., Witte, M. G., van der Heide, U. A., Pos, F. J., Lebesque, J. V., van Herk, M., Incrocci, L., and Heemsbergen, W. D. (2015). Dose–surface maps identifying local dose–effects for acute gastrointestinal toxicity after radiotherapy for prostate cancer. *Radiotherapy and Oncology*, 117(3):515 – 520.

[172] Yan, D. (2010). Adaptive Radiotherapy: Merging Principle Into Clinical Practice. *Seminars in Radiation Oncology*, 20(2):79–83.

[173] Yan, D., Jaffray, D. A., and Wong, J. W. (1999). A model to accumulate fractionated dose in a deforming organ. *International Journal of Radiation Oncology Biology Physics*, 44(3):665–675.

[174] Yartsev, S., Kron, T., and Van Dyk, J. (2007). Tomotherapy as a tool in image-guided radiation therapy (IGRT): theoretical and technological aspects. *Biomedical imaging and intervention journal*, 3(1):e16.

[175] Zambrano, V., Furtado, H., Fabri, D., LUtgendorf-Caucig, C., GOra, J., Stock, M., Mayer, R., Birkfellner, W., and Georg, D. (2013). Performance validation of deformable image registration in the pelvic region. *Journal of Radiation Research*, 54(suppl 1):i120–i128.

[176] Zhao, B., Joiner, M. C., Orton, C. G., and Burmeister, J. (2010). "saber": A new software tool for radiotherapy treatment plan evaluation. *Medical Physics*, 37(11):5586–5592.

[177] Zhen, X., Chen, J., Zhong, Z., Hrycushko, B., Zhou, L., Jiang, S., Albuquerque, K., and Gu, X. (2017). Deep convolutional neural network with transfer learning for rectum toxicity prediction in cervical cancer radiotherapy: a feasibility study. *Phys. Med. Biol*, 62(21):8246–8263.

Appendix A

Statement of originality

The research conducted within this thesis was only possible due to the multidisciplinary nature of the VoxTox research programme. Much of the analysis was dependent upon systems that were previously developed within VoxTox. The work presented in this thesis is my own, unless stated otherwise. Every attempt has been made to appropriately accredit the scientific contributions of others throughout the thesis; a concise summary is included below for clarity.

Image archiving and patient tokenisation

Retrieval and processing of patient data from the TomoTherapy planning system was implemented by Dr Marina Romanchikova (Medical Physics), and Karl Harrison (Cavendish Laboratory), [132]. Referred to in chapters: 3-6

CheckTomo dose calculation software

The dose calculation algorithm for independent calculation of dose based on MVCT image guidance scans was developed by Dr Simon Thomas, Head of Medical Physics & Clinical Engineering at Addenbrooke's Hospital, who is co-supervisor of this PhD research. Referred to in chapters: 3-6

Determining an alpha-beta ratio for the rectum

In order to combine patient datasets from the 60 Gy in 20, and 74 Gy in 37 fractionation schedules, the 60 Gy group were converted to the equivalent dose in 37 fractions. Our data were fit to the CHHiP constraints, and the α/β ratio was iteratively optimised until the equivalent uniform dose and biologically effective dose were equal between the two groups.

The original calculation spreadsheet was generated by Dr Simon Thomas. My involvement was the independent checking and verifying of the calculation. Referred to in chapters: 3-4, 6-7

Autosegmentation

The autosegmentation software was developed by Prof Michael Sutcliffe, Cambridge University Department of Engineering, and co-supervisor of this PhD thesis. As a consequence of work conducted in this PhD project, several areas of improvement were identified from the original version of the autosegmentation software. Example cases were identified where contours suffered due to restrictions of the region of interest used as the search area for the rectal contour, as well cases where air cavities were not being correctly identified. As a result of these investigations, improvements were implemented to the algorithm by Prof Sutcliffe. The version used for discovery cohort analysis was v1.4, and for consolidation cohort analysis, v1.6. Details of version 1.6 were presented in the article 'Autosegmentation of the rectum on megavoltage image guidance scans' published in Biomedical Physics & Engineering Express in January 2019, co-authored by Prof Sutcliffe and myself [143]. Referred to in chapters: 3-6

Toxicity recording and processing

Electronic case reporting forms were used to interrogate and record patient toxicity. These were designed by Amy Bates, Lead Research Radiographer at Addenbrooke's Hospital, and Dr Jessica Scaife, former research fellow in the VoxTox research programme. Mapping rules to translate raw data to standard toxicity reporting systems were developed by Dr Scaife and Dr David Noble, VoxTox research fellow. Mapping rules were externally verified by Dr Agnieszka Lemanska, University of Surrey. Toxicity spreadsheets were populated and maintained by Amy Bates and Karl Harrison. Referred to in chapters: 3-4, 6-7

Dose surface maps and analysis

The MATLAB code to generate dose surface maps (DSMs) was developed by Dr Simon Thomas, following approaches in the literature. Code for calculation of EUD and dose-widths were also written by Dr Thomas. The code was embedded within the code developed within this thesis in order to process and output results for the patients investigated. Referred to in chapters: 3-7

Discovery cohort analysis

The 109 patients analysed from the discovery cohort were selected according to the availability of baseline data. This was acquired manually as discovery patients were recruited prior to the formalised recording of baseline data. Dr Jessica Scaife retrospectively assessed patient notes in order to determine the status of baseline symptoms. Dr Scaife performed a dose-toxicity analysis using DVH data for these 109 patients [137]. Referred to in chapters: 3,4

Toxicity data for the consolidation cohort

The patient toxicity spreadsheet was frozen in February 2018 for analysis of patients in the consolidation cohort. Freezing of the database was performed by Amy Bates. Processing of raw data into final clinical toxicity scores was overseen by Karl Harrison. Data processing to finalise the list of patients with a minimum of 2 year follow up was performed by Dr David Noble. I performed further processing of patient toxicity data following these stages in order to split by staging, interpret baselines and patient characteristics, record dose prescription, then combine all toxicity, baseline, and dose information. The final list of patients was then split into training and validation sets by toxicity rate and prescription. Referred to in chapters: 3-4, 6-7

Finite element model

The biomechanical finite element model was initially developed by Prof Michael Sutcliffe for 'proof of concept' experimental measurements. Prof Sutcliffe developed the MATLAB code for writing the input files to be run in Abaqus, and Python code for extracting final model coordinates. My involvement was: modifying the code so that it could read from patient DICOM files; optimise the element type and hyperelastic properties to represent the anatomy; and introduce an additional function in MATLAB to calculate the dose based on the Abaqus output. In addition to parameterising the resulting FE-DSMs using EUD and dose-widths, I developed the code for division of the rectal wall into geometric subregions, and probabilistic subregions at risk. Referred to in chapters: 3, 5-7

The motion-dose sensitivity study conducted in Chapter 4, and the NTCP modelling conducted in Chapter 7 are my own work.

Appendix B

Research outcomes

First author publications

Shelley, L.E.A & Sutcliffe, M.P.F., Harrison, K., Scaife, J.E., Parker, M.A., Romanchikova, M., Thomas, S.J., Jena, R., Burnet, N.G., Autosegmentation of the rectum on megavoltage image guidance scans, Biomed. Phys. Eng. Express, 5, 025006, 2019 (Figure B.1a).

Shelley, L.E.A, Scaife, J.E., Romanchikova, M., Harrison, K., Forman, J.R., Bates, A.M, Noble, D.J., Jena, R., Parker, M.A., Sutcliffe, M.P.F., Thomas, S.J., Burnet, N. G., Delivered dose can be a better predictor of rectal toxicity than planned dose in prostate radiotherapy, Radiotherapy and Oncology, 123(3), 466-471, 2017 (Figure B.1b).

Contributing author publications

Noble, D.J., Yeap, P.L., Seah, S.Y.K., Harrison, K., **Shelley, L.E.A.**, Romanchikova, M., Bates, A.M., Zheng, Y., Barnett, G.C., Benson, R.J., Jefferies, S.J., Thomas, S.J., Jena, R., Burnet, N.G., Anatomical change during radiotherapy for head and neck cancer, and its effect on delivered dose to the spinal cord. Radiother Oncol, 130. pp. 32-38, 2018.

Neil G Burnet, Jessica E Scaife, Marina Romanchikova, Simon J Thomas, Amy M Bates, Emma Wong, David J Noble, **Leila EA Shelley**, Simon J Bond, Julia R Forman, Andrew CF Hoole, Gillian C Barnett, Frederic M Brochu, Michael PD Simmons, Raj Jena, Karl Harrison, Ping Lin Yeap, Amelia Drew, Emma Silvester, Patrick Elwood, Hannah Pullen, Andrew Sultana, Shannon YK Seah, Megan Z Wilson, Simon G Russell, Richard J Benson, Yvonne L Rimmer, Sarah J Jefferies, Nicolette Taku, Mark Gurnell, Andrew S Powlson, Carola-Bibiane

Schönlieb, Xiaohao Cai, Michael PF Sutcliffe, and Michael A Parker, Applying physical science techniques and CERN technology to an unsolved problem in radiation treatment for cancer: the multidisciplinary 'VoxTox' research programme, CERN Ideasq J Exp Innov.; 1(1): 3–12, 2017.

Published abstracts

Shelley, L., SP-0040 Exploiting IGRT to calculate delivered dose for normal tissue sparing, Radiotherapy and Oncology 133:S12-S13, 2019.

Shelley, L.E.A, Noble, D.J., Romanchikova, M., Harrison, K., Bates, A.M., Sutcliffe, M.P.F., Thomas, S.J., Burnet, N.G., Jena, R., PO-0983 Accumulating delivered dose to the rectum using finite element analysis improves toxicity prediction, Radiotherapy and Oncology 133:S538-S539, 2019.

Noble, D.J., Harrison, K., Hoole, A., Wilson, Z.M., Thomas, S.J., Bates, A.M., **Shelley, L.E.A**, Burnet, N.G., Jena, R., PO-0984 Univariate toxicity associations are stronger with delivered than planned dose in HNC patients, Radiotherapy and Oncology 133:S539-S540, 2019.

Niebuhr, N.I., Bostel, T., Harrison, K., Jena, R., Nicolay, N.H., Noble, D.J., **Shelley, L.E.A**, Splinter, M., Pfaffenberger, A., EP-1922 Comparing biological and conventional dose accumulation using daily imaging of head and neck and pelvis cases, Radiotherapy and Oncology 133:S1045-S1046, 2019.

Shelley, L.E.A, Noble, D.J., Romanchikova, M., Harrison, K., Bates, A.M., Sutcliffe, M.P.F., Thomas, S.J., Burnet, N.G., EP-2010: NTCP model for rectal bleeding based on delivered dose surface maps in prostate radiotherapy, Radiotherapy and Oncology 127:S1095-S1096, 2018.

Noble, D.J., Harrison, K., Yeap, P.L., Seah, S.Y.K., Zheng, Y., Bates, A.M., Romanchikova, M., **Shelley, L.E.A.**, Thomas, S.J., Parker, M.A., Burnet, N.G., PO-0961: Anatomical predictors of differences between planned and delivered dose to Head and neck OARs, Radiotherapy and Oncology 127:S527-S528, 2018.

Shelley, L.E.A, Scaife, J.E., Bates, A.M, Forman, J.R., Harrison, K., Jena, R., Noble, D.J., Parker, M.A., Romanchikova, M., Sutcliffe, M.P.F., Thomas, S.J., Burnet, N. G., PO-0880: Using accumulated delivered dose to predict rectal toxicity in prostate radiotherapy, Radiotherapy and Oncology 123:S481-S482, 2017.

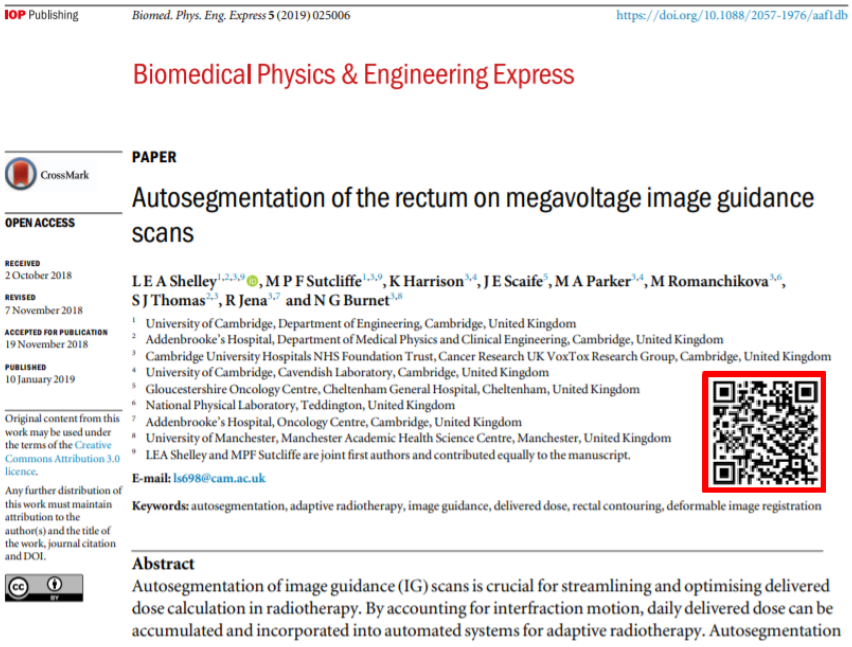
Invited speaker

ESTRO 38, Milan, 26th April 2019, Exploiting IGRT to calculate delivered dose for normal tissue sparing.

Scottish Radiotherapy Research Forum (ScoRRF), Stirling, 2nd November 2017, The VoxTox Project: Using delivered dose to improve dose-toxicity correlation.

Other

- Member of Institute of Physics and Engineering in Medicine (IPEM) Radiotherapy Special Interest Group, 2014-2016
- Member of the General Council of European Alliance for Medical and Biological Engineering & Science (EAMBES), 2016-2018
- Co-supervision of MEng student, Carla Troyas, on use of machine learning for autosegmentation of the rectum.
- STEM outreach activities including the Cavendish Laboratory 'Physics at work' day, and CUED bioengineering open day.
- Interview for 'She Loves MATLAB', https://blogs.mathworks.com/community/2018



IOP Publishing Biomed. Phys. Eng. Express 5 (2019) 025006 https://doi.org/10.1088/2057-1976/aaf1db

Biomedical Physics & Engineering Express

PAPER

Autosegmentation of the rectum on megavoltage image guidance scans

L E A Shelley[1,2,3,9], M P F Sutcliffe[1,3,9], K Harrison[3,4], J E Scaife[5], M A Parker[3,4], M Romanchikova[3,6], S J Thomas[2,3], R Jena[3,7] and N G Burnet[3,8]

[1] University of Cambridge, Department of Engineering, Cambridge, United Kingdom
[2] Addenbrooke's Hospital, Department of Medical Physics and Clinical Engineering, Cambridge, United Kingdom
[3] Cambridge University Hospitals NHS Foundation Trust, Cancer Research UK VoxTox Research Group, Cambridge, United Kingdom
[4] University of Cambridge, Cavendish Laboratory, Cambridge, United Kingdom
[5] Gloucestershire Oncology Centre, Cheltenham General Hospital, Cheltenham, United Kingdom
[6] National Physical Laboratory, Teddington, United Kingdom
[7] Addenbrooke's Hospital, Oncology Centre, Cambridge, United Kingdom
[8] University of Manchester, Manchester Academic Health Science Centre, Manchester, United Kingdom
[9] LEA Shelley and MPF Sutcliffe are joint first authors and contributed equally to the manuscript.

E-mail: ls698@cam.ac.uk

Keywords: autosegmentation, adaptive radiotherapy, image guidance, delivered dose, rectal contouring, deformable image registration

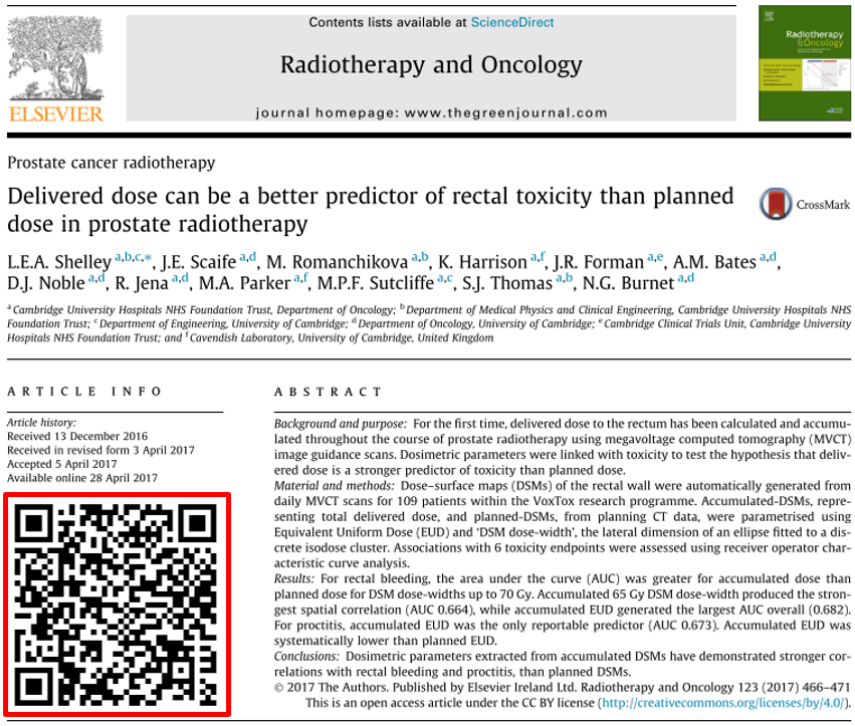
OPEN ACCESS

RECEIVED 2 October 2018
REVISED 7 November 2018
ACCEPTED FOR PUBLICATION 19 November 2018
PUBLISHED 10 January 2019

Abstract

Autosegmentation of image guidance (IG) scans is crucial for streamlining and optimising delivered dose calculation in radiotherapy. By accounting for interfraction motion, daily delivered dose can be accumulated and incorporated into automated systems for adaptive radiotherapy. Autosegmentation

(a)



Contents lists available at ScienceDirect

Radiotherapy and Oncology

journal homepage: www.thegreenjournal.com

Prostate cancer radiotherapy

Delivered dose can be a better predictor of rectal toxicity than planned dose in prostate radiotherapy

L.E.A. Shelley[a,b,c,*], J.E. Scaife[a,d], M. Romanchikova[a,b], K. Harrison[a,f], J.R. Forman[a,e], A.M. Bates[a,d], D.J. Noble[a,d], R. Jena[a,d], M.A. Parker[a,f], M.P.F. Sutcliffe[a,c], S.J. Thomas[a,b], N.G. Burnet[a,d]

[a] Cambridge University Hospitals NHS Foundation Trust, Department of Oncology; [b] Department of Medical Physics and Clinical Engineering, Cambridge University Hospitals NHS Foundation Trust; [c] Department of Engineering, University of Cambridge; [d] Department of Oncology, University of Cambridge; [e] Cambridge Clinical Trials Unit, Cambridge University Hospitals NHS Foundation Trust; and [f] Cavendish Laboratory, University of Cambridge, United Kingdom

ARTICLE INFO

Article history:
Received 13 December 2016
Received in revised form 3 April 2017
Accepted 5 April 2017
Available online 28 April 2017

ABSTRACT

Background and purpose: For the first time, delivered dose to the rectum has been calculated and accumulated throughout the course of prostate radiotherapy using megavoltage computed tomography (MVCT) image guidance scans. Dosimetric parameters were linked with toxicity to test the hypothesis that delivered dose is a stronger predictor of toxicity than planned dose.

Material and methods: Dose–surface maps (DSMs) of the rectal wall were automatically generated from daily MVCT scans for 109 patients within the VoxTox research programme. Accumulated-DSMs, representing total delivered dose, and planned-DSMs, from planning CT data, were parametrised using Equivalent Uniform Dose (EUD) and 'DSM dose-width', the lateral dimension of an ellipse fitted to a discrete isodose cluster. Associations with 6 toxicity endpoints were assessed using receiver operator characteristic curve analysis.

Results: For rectal bleeding, the area under the curve (AUC) was greater for accumulated dose than planned dose for DSM dose-widths up to 70 Gy. Accumulated 65 Gy DSM dose-width produced the strongest spatial correlation (AUC 0.664), while accumulated EUD generated the largest AUC overall (0.682). For proctitis, accumulated EUD was the only reportable predictor (AUC 0.673). Accumulated EUD was systematically lower than planned EUD.

Conclusions: Dosimetric parameters extracted from accumulated DSMs have demonstrated stronger correlations with rectal bleeding and proctitis, than planned DSMs.

(b)

Fig. B.1 First author publications in (a) Biomedical Physics and Engineering Express, 2019, and (b) Radiotherapy and Oncology, 2017. Use the camera on a smartphone or tablet to focus on the QR codes in the red boxes to link to articles. A QR reader may need to be installed.